GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 30, 2002, 07:12:43 ; Search time 2325.7 Seconds
(without alignments)
992.189 Million cell updates/sec

Title: US-09-303-518D-127
Perfect score: 1344
Sequence: 1 atgattaaaatcaaaaaagg.....ccnttgagaaggaaggctga 1344

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1327	98.7	1344	20	AAZ12027	Neisseria meningit
	2	1229.4	91.5	1344	20	AAZ12026	Neisseria meningit
c	3	1229.4	91.5	44608	21	AAA81495	N. meningitidis pa
c	4	1229.4	91.5	349980	21	AAF21607	Neisseria meningit
	5	1165.4	86.7	1344	20	AAZ12028	Neisseria gonorrho
	6	527.2	39.2	1830121	17	AAT42063	Haemophilus influe
	7	431.6	32.1	474	20	AAZ12025	Neisseria meningit
	8	431.6	32.1	474	21	AAA81335	N. meningitidis Me
c	9	351	26.1	363	21	AAZ54035	Neisseria meningit
c	10	333.6	24.8	96109	22	AAF28548	Genomic fragment #
c	11	317.4	23.6	363	21	AAZ54034	Neisseria meningit
c	12	291.8	21.7	363	21	AAZ54033	Neisseria gonorrhe
	13	122.2	9.1	1353	20	AAX91657	Porphorymonas ging
	14	122.2	9.1	1362	20	AAX91536	Porphorymonas ging
c	15	39.2	2.9	5059	20	AAX84332	Stealth virus nucl
	16	38.8	2.9	9762	21	AAZ48249	Rubella virus Cend
	17	38.2	2.8	65140	22	AAD17184	Streptomyces nours
	18	38.2	2.8	125401	22	AAD17186	Streptomyces nours
	19	37.2	2.8	9759	19	AAV34766	Rubella virus RA27
c	20	37	2.8	763	20	AAZ09101	Mistletoe lectin A
c	21	37	2.8	1598	20	AAZ09100	Mistletoe lectin D
	22	36.6	2.7	1401	21	AAZ95378	Chlamydia pneumoni
	23	36.4	2.7	1325	22	AAD08837	Human G-protein co
c	24	36.4	2.7	4096	22	AAF87442	Corynebacterium th
c	25	36	2.7	762	20	AAZ09104	Mistletoe lectin A
	26	36	2.7	3230	22	AAH16235	Human cDNA sequenc
c	27	36	2.7	92934	21	AAA81473	N. meningitidis pa
c	28	36	2.7	172325	21	AAF21613	Neisseria meningit
c	29	36	2.7	837096	21	AAA81489	N. meningitidis pa
	30	35.8	2.7	699	22	AAC97005	Mycobacterium tube
c	31	35.8	2.7	4403765	22	AAI99683	Mycobacterium tube
c	32	35.8	2.7	4411529	22	AAI99682	Mycobacterium tube
	33	35.6	2.6	2558	21	AAA08045	f-Therien rubella
	34	35.6	2.6	9757	16	AAQ97686	Infectious rubella
	35	35.6	2.6	9759	18	AAT89642	Infectious rubella
	36	35.6	2.6	9759	21	AAA08043	Rubella virus cDNA
	37	35.4	2.6	2727	23	AAS54113	Pseudomonas aerugi
	38	35	2.6	296	23	AAS57519	cDNA #195 encoding
	39	35	2.6	1935	19	AAV23811	Nitrosomonas dnaK
	40	34.6	2.6	323	23	AAS86095	DNA encoding novel
	41	34.6	2.6	1139	21	AAF21735	Human breast and o
	42	34.6	2.6	1689	23	AAS86098	DNA encoding novel
	43	34.6	2.6	6237	23	AAS76357	DNA encoding novel
c	44	34.4	2.6	1050	23	AAS73416	DNA encoding novel
c	45	34.4	2.6	1050	23	AAS80144	DNA encoding novel
	46	34.4	2.6	1774	22	AAK52900	Human polynucleoti
	47	34.4	2.6	2165	18	AAT59702	Endotoxin CryV gen
c	48	34.4	2.6	18974	21	AAA81485	N. meningitidis pa
	49	34.4	2.6	349980	21	AAF21607	Neisseria meningit
	50	34.4	2.6	1437668	21	AAA81490	N. meningitidis B
	51	34.2	2.5	1116	22	AAF28683	Human protein HP03
	52	34.2	2.5	1116	24	ABA02176	Human G protein-co
	53	34.2	2.5	1119	21	AAA64367	DNA encoding a P2Y
	54	34.2	2.5	1119	21	AAD01119	Human orphan G pro
	55	34.2	2.5	1119	21	AAA46018	Human G protein co
	56	34.2	2.5	1119	22	AAH49504	Human GTP-binding
	57	34.2	2.5	1119	22	AAD02585	Human G-protein co
	58	34.2	2.5	1119	22	AAF86237	Human G-protein co
	59	34.2	2.5	1119	24	AAS98045	Human DNA for pote
c	60	34.2	2.5	1237	24	AAS98085	Human DNA for pote
	61	34.2	2.5	1560	24	AAS19414	Human cDNA encodin
	62	34.2	2.5	1720	22	AAF28693	Human protein HP03
	63	34.2	2.5	2444	24	AAD26369	Human G-protein co
	64	34.2	2.5	2480	22	AAD06509	Human CON217 G pro
	65	34.2	2.5	2559	21	AAZ95039	Human G protein co
	66	34.2	2.5	3180	22	AAF25830	Human P2YLi DNA.
c	67	34.2	2.5	3789	22	AAH66883	C glutamicum codin
	68	34.2	2.5	15783	22	AAS39803	Genomic sequence #
	69	34.2	2.5	15783	22	AAK90159	Human digestive sy
c	70	34.2	2.5	25365	23	AAS59558	Propionibacterium
c	71	34.2	2.5	349980	22	AAH68529	C glutamicum codin
c	72	34.2	2.5	349980	22	AAH68530	C glutamicum codin
	73	34	2.5	10577	21	AAA81494	N. meningitidis pa
	74	34	2.5	44377	18	AAT78508	Platenolide syntha
	75	34	2.5	44377	18	AAT80414	Platenolide syntha
	76	34	2.5	69936	21	AAA81479	N. meningitidis pa
	77	34	2.5	349980	21	AAF21610	Neisseria meningit
c	78	33.8	2.5	18796	23	AAS59517	Propionibacterium
	79	33.8	2.5	29379	23	AAS59510	Propionibacterium
	80	33.6	2.5	621	22	AAF26372	P. putida oxygenas
	81	33.6	2.5	1662	20	AAX86154	DNA encoding a Str
	82	33.6	2.5	10732	21	AAA10594	Gene encoding a su

83	33.4	2.5	4590	22	AAH24065	Yeast AOD9604-asso
84	33.2	2.5	1544	23	ABL06725	Drosophila melanog
85	33.2	2.5	3866	23	ABL06724	Drosophila melanog
c 86	33	2.5	245	21	AAA44957	Human secreted exp
c 87	33	2.5	1311	21	AAC42295	Arabidopsis thalia
88	33	2.5	5416	19	AAV07381	Neisseria meningit
89	32.8	2.4	876	23	ABL21113	Drosophila melanog
90	32.8	2.4	1137	23	ABL16599	Drosophila melanog
c 91	32.8	2.4	1596	20	AAZ09103	Mistletoe lectin I
c 92	32.8	2.4	4399	23	ABL16598	Drosophila melanog
c 93	32.8	2.4	4405	23	ABL21112	Drosophila melanog
94	32.6	2.4	1451	21	AAF14884	Trichoderma reesei
95	32.6	2.4	14041	22	AAH48024	Internal control B
96	32.4	2.4	754	22	AAH00214	Stenotrophomonas m
c 97	32.4	2.4	1191	22	AAH66975	C glutamicum codin
98	32.4	2.4	1400	13	AAQ22689	Sequence encoding
c 99	32.4	2.4	2394	21	AAA15156	DNA encoding outer
c 100	32.4	2.4	2394	22	AAS07279	Neisseria gonorrho

ALIGNMENTS

RESULT 1
AAZ12027
ID AAZ12027 standard; DNA; 1344 BP.
XX
AC AAZ12027;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria meningitidis strain A complete ORF22 sequence.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS Neisseria meningitidis.
XX
PN WO9924578-A2.
XX
PD 20-MAY-1999.
XX
PF 09-OCT-1998; 98WO-IB01665.
XX
PR 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX
DR WPI; 1999-327407/27.
DR P-PSDB; AAY38562.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
PS Claim 9; Page 123; 524pp; English.
XX
CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicaemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX

SQ Sequence 1344 BP; 323 A; 347 C; 368 G; 289 T; 17 other;

Query Match 98.7%; Score 1327; DB 20; Length 1344;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcatt 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcatt 60

Qy     61 tatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Qy    121 ccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccat 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 aaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccat 240

Qy    241 cgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 cgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatc 300

Qy    301 gagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngc 360

Qy    361 aatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccct 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 aatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccct 420

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctngcg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctngcg 480

Qy    481 gcagaccctgtggttgtgatcaaagaagccgncgangatttcagacgangtntgctggta 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 gcagaccctgtggttgtgatcaaagaagccgncgangatttcagacgangtntgctggta 540

Qy    541 ttgagccgtttgaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgccg 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ttgagccgtttgaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgccg 600

Qy    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatccggccggtttg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatccggccggtttg 660

Qy    661 agtggcacgcacattcatttcattgagccggtcggtgcaaacaaaaccgtttggaccatc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 agtggcacgcacattcatttcattgagccggtcggtgcaaacaaaaccgtttggaccatc 720

Qy    721 aattatcaagatgtaattgccatcggacgtttgtttgcaacaggccgtctgaacaccgag 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 aattatcaagatgtaattgccatcggacgtttgtttgcaacaggccgtctgaacaccgag 780

Qy    781 cgcgtgattgctttgggtggttctcaagtcaacaaaccacgcctcttgcgtaccgttttg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 cgcgtgattgctttgggtggttctcaagtcaacaaaccacgcctcttgcgtaccgttttg 840

Qy    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacgcagacaaccgcgtgatt 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacgcagacaaccgcgtgatt 900

Qy    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020

Qy   1021 ccgcagccggacaaatactccatcacgcgtacgaccctcggccatttcctgaaaaacaaa 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ccgcagccggacaaatactccatcacgcgtacgaccctcggccatttcctgaaaaacaaa 1080

Qy   1081 ctcttcaagttcacgacagccgtcaacggtggcgaccgcgccatggtgccgattggtact 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ctcttcaagttcacgacagccgtcaacggtggcgaccgcgccatggtgccgattggtact 1140

Qy   1141 tacgagcgcgtaatgccgctagacatcctgcctaccctgcttttgcgcgatttaatcgtc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 tacgagcgcgtaatgccgctagacatcctgcctaccctgcttttgcgcgatttaatcgtc 1200

Qy   1201 ggcgataccgacagcgcgcaagcattgggttgcttggaattggacgaagaagacctcgct 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ggcgataccgacagcgcgcaagcattgggttgcttggaattggacgaagaagacctcgct 1260

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatanggcccgctgttgcgtaaggtgctg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatanggcccgctgttgcgtaaggtgctg 1320

Qy   1321 gaaaccnttgagaaggaaggctga 1344
          |||||||||||||||||||||||
Db   1321 gaaaccnttgagaaggaaggctga 1344
```

RESULT 2
AAZ12026
ID AAZ12026 standard; DNA; 1344 BP.
XX
AC AAZ12026;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria meningitidis complete ORF22 sequence.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS Neisseria meningitidis.
XX
PN WO9924578-A2.
XX
PD 20-MAY-1999.
XX
PF 09-OCT-1998; 98WO-IB01665.
XX
PR 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX
DR WPI; 1999-327407/27.
DR P-PSDB; AAY38561.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
PS Claim 9; Page 122-123; 524pp; English.
XX
CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicaemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX
SQ Sequence 1344 BP; 329 A; 366 C; 365 G; 284 T; 0 other;

Query Match 91.5%; Score 1229.4; DB 20; Length 1344;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 1266; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

```
Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcatt 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| | ||
Db      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60

Qy     61 tatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
          || ||||| || | |||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Qy    121 ccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccat 240
          |||||| |||||||||||||||||| ||||| | |||||||||||||||||| || ||
Db    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240

Qy    241 cgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatc 300
          || |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300

Qy    301 gagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngc 360
          ||||| |||||||||||| || |||||| |||||||||||||||||||| | | | ||
Db    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360

Qy    361 aatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccct 420
          || |||||||||||||||||||||||||||||||| | |||||||||||||||||| |||
Db    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctngcg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480

Qy    481 gcagaccctgtggttgtgatcaaagaagccgncgangatttcagacgangtntgctggta 540
          || ||||||  |||  | |||||||||||||| ||| ||||||| |||  |  || |||||
Db    481 gccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttggta 540

Qy    541 ttgagccgtttgaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgccg 600
          |||||||||||||||| || |||||||||| ||||||||||||||||||||||||||||||
Db    541 ttgagccgtttgaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgccg 600

Qy    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatccggccggtttg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtttg 660

Qy    661 agtggcacgcacattcatttcattgagccggtcggtgcaaacaaaaccgtttggaccatc 720
          |||||||||||||||||||||| |||||||||| || || |||||||| |||||||||
Db    661 agtggcacgcacattcatttcatcgagccggtcggcgcgaataaaaccgtgtggaccatc 720

Qy    721 aattatcaagatgtaattgccatcggacgtttgtttgcaacaggccgtctgaacaccgag 780
          |||||||||||||||||| |||| || ||||||||||||||||||||||||||||||||||
Db    721 aattatcaagatgtaattaccattggccgtttgtttgcaacaggccgtctgaacaccgag 780

Qy    781 cgcgtgattgctttgggtggttctcaagtcaacaaaccacgcctcttgcgtaccgttttg 840
          |||||||||||  | |||||||||||||||||||||||| ||||||||||||||||||||
Db    781 cgcgtgattgccctaggtggttctcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840

Qy    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacgcagacaaccgcgtgatt 900
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacacagacaaccgcgtgatt 900

Qy    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020

Qy   1021 ccgcagccggacaaatactccatcacgcgtacgaccctcggccatttcctgaaaaacaaa 1080
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1021 ccgcagccggacaaatactccatcacgcgtacaaccctcggccatttcctgaaaaacaaa 1080

Qy   1081 ctcttcaagttcacgacagccgtcaacggtggcgaccgcgccatggtgccgattggtact 1140
          |||||||||||||  |||||||||||||| ||||||||||||||||||||||||||||||
Db   1081 ctcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggtgccgattggtact 1140

Qy   1141 tacgagcgcgtaatgccgctagacatcctgcctaccctgcttttgcgcgatttaatcgtc 1200
          |||||||||||| |||||  | || |||||||| ||||||||||||||||||||||||||||
Db   1141 tacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcgcgatttaatcgtc 1200

Qy   1201 ggcgataccgacagcgcgcaagcattgggttgcttggaattggacgaagaagacctcgct 1260
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   1201 ggcgataccgacagcgcgcaggcattgggttgcttggaattggacgaagaagacctcgct 1260

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatanggcccgctgttgcgtaaggtgctg 1320
          |||||||||||||||||||||||||||||||| |||||||||||| || ||||||
Db   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320

Qy   1321 gaaaccnttgagaaggaaggctga 1344
          |||||| ||||||||||||||||||
Db   1321 gaaaccattgagaaggaaggctga 1344
```

RESULT 3
AAA81495/c
ID AAA81495 standard; DNA; 44608 BP.
XX
AC AAA81495;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_42 SEQ ID NO:42.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Claim 7; Page 1283-1296; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neissariae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
SQ Sequence 44608 BP; 10938 A; 10835 C; 11999 G; 10834 T; 2 other;

Query Match 91.5%; Score 1229.4; DB 21; Length 44608;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 1266; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

```
Qy       1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcatt 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| | ||
Db   25944 ATGATTAAAATCAAAAAAGGTCTAAACCTGCCCATCGCGGGCAGACCGGAGCAAGCCGTT 25885

Qy      61 tatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
           || ||||| || | ||||||||||||||||||||||||||||||||||||||||||||||
Db   25884 TACGACGGCCCGGCCATTACCGAAGTCGCGTTGCTTGGCGAAGAATATGCCGGTATGCGC 25825

Qy     121 ccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25824 CCCTCGATGAAAGTCAAGGAAGGCGATGCCGTCAAAAAAGGCCAAGTGCTGTTTGAAGAC 25765

Qy     181 aaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccat 240
           |||||| |||||||||||||||||| ||||| | |||||||||||||||| || ||
Db   25764 AAAAAGAATCCGGGCGTGGTGTTTACTGCGCCGGCTTCAGGCAAAATCGCCGCGATTCAC 25705

Qy     241 cgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatc 300
           || |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   25704 CGTGGCGAAAAGCGCGTACTTCAGTCAGTCGTGATTGCCGTTGAAGGCAACGACGAAATC 25645

Qy     301 gagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngc 360
           ||||| |||||||||| || |||||| |||||||||||||||||| ||| | | ||
Db   25644 GAGTTTGAACGCTACGCACCTGAAGCGCTGGCAAACTTAAGCGGCGAAGAAGTGCGCCGC 25585

Qy     361 aatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccct 420
           || |||||||||||||||||||||||||||||| | |||||||||||||||||| |||
Db   25584 AACCTGATCCAATCCGGTTTGTGGACTGCGCTGCGCACCCGTCCGTTCAGCAAAATTCCT 25525

Qy     421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctngcg 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   25524 GCCGTCGATGCCGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCT 25465

Qy     481 gcagaccctgtggttgtgatcaaagaagccgncgangatttcagacgangtntgctggta 540
           || |||||| ||| | |||||||||||||| ||| ||||||| ||| | || |||||
Db   25464 GCCGACCCTACGGTCATTATCAAAGAAGCCGCCGAGGATTTCAAACGCGGCCTGTTGGTA 25405

Qy     541 ttgagccgtttgaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgccg 600
           |||||||||||||||| || |||||||||| |||||||||||||||||||||||||||||
Db   25404 TTGAGCCGTTTGACCGAACGCAAAATCCATGTTTGTAAGGCAGCTGGCGCAGACGTGCCG 25345

Qy     601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatccggccggtttg 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   25344 TCTGAAAATGCTGCCAACATCGAAACACATGAATTCGGCGGCCCGCATCCTGCCGGTTTG 25285
```

```
Qy    661 agtggcacgcacattcatttcattgagccggtcggtgcaaacaaaaccgtttggaccatc 720
          |||||||||||||||||||||||| |||||||||| || || |||||||| |||||||||
Db  25284 AGTGGCACGCACATTCATTTCATCGAGCCGGTCGGCGCGAATAAAACCGTGTGGACCATC 25225

Qy    721 aattatcaagatgtaattgccatcggacgtttgtttgcaacaggccgtctgaacaccgag 780
          ||||||||||||||||| |||| || ||||||||||||||||||||||||||||||||||
Db  25224 AATTATCAAGATGTAATTACCATTGGCCGTTTGTTTGCAACAGGCCGTCTGAACACCGAG 25165

Qy    781 cgcgtgattgctttgggtggttctcaagtcaacaaaccacgcctcttgcgtaccgttttg 840
          |||||||||||  | ||||||||||||||||||||| ||||||||||||||||||||
Db  25164 CGCGTGATTGCCCTAGGTGGTTCTCAAGTCAACAAACCGCGCCTCTTGCGTACCGTTTTG 25105

Qy    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacgcagacaaccgcgtgatt 900
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  25104 GGTGCGAAAGTATCGCAAATTACTGCGGGCGAATTGGTTGACACAGACAACCGCGTGATT 25045

Qy    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  25044 TCCGGTTCGGTATTGAACGGCGCGATTACACAAGGCGCGCACGATTATTTGGGACGCTAC 24985

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  24984 CACAATCAGATTTCCGTTATCGAAGAAGGCCGCAGCAAAGAGCTGTTCGGCTGGGTTGCG 24925

Qy   1021 ccgcagccggacaaatactccatcacgcgtacgaccctcggccatttcctgaaaaacaaa 1080
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  24924 CCGCAGCCGGACAAATACTCCATCACGCGTACAACCCTCGGCCATTTCCTGAAAAACAAA 24865

Qy   1081 ctcttcaagttcacgacagccgtcaacggtggcgaccgcgccatggtgccgattggtact 1140
          |||||||||||| |||||||||||||| ||||||||||||||||||||||||||||||||
Db  24864 CTCTTCAAGTTCAACACAGCCGTCAACGGCGGCGACCGCGCCATGGTGCCGATTGGTACT 24805

Qy   1141 tacgagcgcgtaatgccgctagacatcctgcctaccctgcttttgcgcgatttaatcgtc 1200
          |||||||||| |||||  | || |||||||| ||||||||||||||||||||||||||||
Db  24804 TACGAGCGCGTGATGCCCTTGGATATCCTGCCCACCCTGCTTTTGCGCGATTTAATCGTC 24745

Qy   1201 ggcgataccgacagcgcgcaagcattgggttgcttggaattggacgaagaagacctcgct 1260
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  24744 GGCGATACCGACAGCGCGCAGGCATTGGGTTGCTTGGAATTGGACGAAGAAGACCTCGCT 24685

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatanggcccgctgttgcgtaaggtgctg 1320
          |||||||||||||||||||||||||||||||||||| |||||||||||||| || ||||||
Db  24684 TTGTGCAGCTTCGTCTGCCCGGGCAAATACGAATACGGCCCGCTGTTGCGCAAAGTGCTG 24625

Qy   1321 gaaaccnttgagaaggaaggctga 1344
          |||||| ||||||||||||||||
Db  24624 GAAACCATTGAGAAGGAAGGCTGA 24601
```

RESULT 4
AAF21607/c
ID AAF21607 standard; DNA; 349980 BP.
XX
AC AAF21607;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:108.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
OS Neisseria meningitidis.
XX
PN WO200066791-A1.
XX
PD 09-NOV-2000.
XX
PF 08-MAR-2000; 2000WO-US05928.
XX
PR 30-APR-1999; 99US-0132068.
PR 08-OCT-1999; 99WO-US23573.
PR 28-FEB-2000; 2000GB-0004695.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX
DR WPI; 2000-647603/62.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
XX
PS Claim 7; Appendix A; 692pp; English.
XX
CC The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.
XX
SQ Sequence 349980 BP; 84410 A; 84863 C; 94187 G; 86520 T; 0 other;

Query Match 91.5%; Score 1229.4; DB 21; Length 349980;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 1266; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

```
Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcatt 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| | ||
Db 297909 ATGATTAAAATCAAAAAAGGTCTAAACCTGCCCATCGCGGGCAGACCGGAGCAAGCCGTT 297850

Qy     61 tatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
          || ||||| || | ||||||||||||||||||||||||||||||||||||||||||||||
Db 297849 TACGACGGCCCGGCCATTACCGAAGTCGCGTTGCTTGGCGAAGAATATGCCGGTATGCGC 297790

Qy    121 ccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297789 CCCTCGATGAAAGTCAAGGAAGGCGATGCCGTCAAAAAAGGCCAAGTGCTGTTTGAAGAC 297730

Qy    181 aaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccat 240
          |||||| |||||||||||||||||| ||||| | |||||||||||||||||| || ||
Db 297729 AAAAAGAATCCGGGCGTGGTGTTTACTGCGCCGGCTTCAGGCAAAATCGCCGCGATTCAC 297670

Qy    241 cgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatc 300
          || ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db 297669 CGTGGCGAAAAGCGCGTACTTCAGTCAGTCGTGATTGCCGTTGAAGGCAACGACGAAATC 297610

Qy    301 gagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngc 360
          ||||| |||||||||| || |||||| |||||||||||||||||| ||| |  |  ||
Db 297609 GAGTTTGAACGCTACGCACCTGAAGCGCTGGCAAACTTAAGCGGCGAAGAAGTGCGCCGC 297550

Qy    361 aatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccct 420
          || |||||||||||||||||||||||||||||||| | |||||||||||||||| |||
```

```
Db 297549 AACCTGATCCAATCCGGTTTGTGGACTGCGCTGCGCACCCGTCCGTTCAGCAAAATTCCT 297490

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctngcg 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db 297489 GCCGTCGATGCCGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCT 297430

Qy    481 gcagaccctgtggttgtgatcaaagaagccgncgangatttcagacgangtntgctggta 540
          || |||||| ||| | |||||||||||||| ||| ||||||| ||| | || |||||
Db 297429 GCCGACCCTACGGTCATTATCAAAGAAGCCGCCGAGGATTTCAAACGCGGCCTGTTGGTA 297370

Qy    541 ttgagccgtttgaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgccg 600
          |||||||||||||||| || |||||||||| ||||||||||||||||||||||||||||
Db 297369 TTGAGCCGTTTGACCGAACGCAAAATCCATGTTTGTAAGGCAGCTGGCGCAGACGTGCCG 297310

Qy    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatccggccggtttg 660
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db 297309 TCTGAAAATGCTGCCAACATCGAAACACATGAATTCGGCGGCCCGCATCCTGCCGGTTTG 297250

Qy    661 agtggcacgcacattcatttcattgagccggtcggtgcaaacaaaaccgtttggaccatc 720
          |||||||||||||||||||| |||||||||| || || |||||||| |||||||||||
Db 297249 AGTGGCACGCACATTCATTTCATCGAGCCGGTCGGCGCGAATAAAACCGTGTGGACCATC 297190

Qy    721 aattatcaagatgtaattgccatcggacgtttgtttgcaacaggccgtctgaacaccgag 780
          |||||||||||||||||| |||| || ||||||||||||||||||||||||||||||||
Db 297189 AATTATCAAGATGTAATTACCATTGGCCGTTTGTTTGCAACAGGCCGTCTGAACACCGAG 297130

Qy    781 cgcgtgattgctttgggtggttctcaagtcaacaaaccacgcctcttgcgtaccgttttg 840
          |||||||||||  | ||||||||||||||||||||| ||||||||||||||||||||||
Db 297129 CGCGTGATTGCCCTAGGTGGTTCTCAAGTCAACAAACCGCGCCTCTTGCGTACCGTTTTG 297070

Qy    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacgcagacaaccgcgtgatt 900
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db 297069 GGTGCGAAAGTATCGCAAATTACTGCGGGCGAATTGGTTGACACAGACAACCGCGTGATT 297010

Qy    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297009 TCCGGTTCGGTATTGAACGGCGCGATTACACAAGGCGCGCACGATTATTTGGGACGCTAC 296950

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 296949 CACAATCAGATTTCCGTTATCGAAGAAGGCCGCAGCAAAGAGCTGTTCGGCTGGGTTGCG 296890

Qy   1021 ccgcagccggacaaatactccatcacgcgtacgaccctcggccatttcctgaaaaacaaa 1080
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db 296889 CCGCAGCCGGACAAATACTCCATCACGCGTACAACCCTCGGCCATTTCCTGAAAAACAAA 296830

Qy   1081 ctcttcaagttcacgacagccgtcaacggtggcgaccgcgccatggtgccgattggtact 1140
          |||||||||||||  |||||||||||||| ||||||||||||||||||||||||||||||
Db 296829 CTCTTCAAGTTCAACACAGCCGTCAACGGCGGCGACCGCGCCATGGTGCCGATTGGTACT 296770

Qy   1141 tacgagcgcgtaatgccgctagacatcctgcctaccctgcttttgcgcgatttaatcgtc 1200
          |||||||||| |||||  | || |||||||| ||||||||||||||||||||||||||||
Db 296769 TACGAGCGCGTGATGCCCTTGGATATCCTGCCCACCCTGCTTTTGCGCGATTTAATCGTC 296710

Qy   1201 ggcgataccgacagcgcgcaagcattgggttgcttggaattggacgaagaagacctcgct 1260
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db 296709 GGCGATACCGACAGCGCGCAGGCATTGGGTTGCTTGGAATTGGACGAAGAAGACCTCGCT 296650

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatanggcccgctgttgcgtaaggtgctg 1320
          |||||||||||||||||||||||||||||||||| |||||||||||||| || ||||||
Db 296649 TTGTGCAGCTTCGTCTGCCCGGGCAAATACGAATACGGCCCGCTGTTGCGCAAAGTGCTG 296590

Qy   1321 gaaaccnttgagaaggaaggctga 1344
          |||||| ||||||||||||||||||
Db 296589 GAAACCATTGAGAAGGAAGGCTGA 296566
```

```
RESULT   5
AAZ12028
ID   AAZ12028 standard; DNA; 1344 BP.
XX
AC   AAZ12028;
XX
DT   08-OCT-1999  (first entry)
XX
DE   Neisseria gonorrhoeae  ORF22 polynucleotide sequence.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW   treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS   Neisseria gonorrhoeae.
XX
PN   WO9924578-A2.
XX
PD   20-MAY-1999.
XX
PF   09-OCT-1998;   98WO-IB01665.
XX
PR   01-SEP-1998;   98GB-0019016.
PR   06-NOV-1997;   97GB-0023516.
PR   14-NOV-1997;   97GB-0024190.
PR   18-NOV-1997;   97GB-0024386.
PR   27-NOV-1997;   97GB-0025158.
PR   10-DEC-1997;   97GB-0026147.
PR   14-JAN-1998;   98GB-0000759.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Grandi G,  Masignani V,  Pizza M,  Rappuoli R,  Scarlato V;
XX
DR   WPI; 1999-327407/27.
DR   P-PSDB; AAY38565.
XX
PT   Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT   diagnosis, treatment and prevention of infection
XX
PS   Claim 9; Page 125; 524pp; English.
XX
CC   Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC   (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC   antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC   fragments, their nucleic acids and antibodies are used for diagnosis,
CC   prevention (as vaccines) or treatment of Neisseria infections,
CC   such as meningitis, septicaemia and gonorrhea. Both organisms
CC   are closely related. Fragments of the nucleic acids are useful
CC   as hybridisation probes and antisense reagents.
XX
SQ   Sequence 1344 BP; 334 A; 363 C; 362 G; 285 T; 0 other;
```

```
  Query Match             86.7%;  Score 1165.4;  DB 20;  Length 1344;
  Best Local Similarity   91.2%;  Pred. No. 0;
  Matches 1226;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcatt 60
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      1 atgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcatt 60

Qy     61 tatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
          |||||||| || | |||||||||||||||||||||||||||||||||| ||| ||||||
Db     61 tatgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgtcggcatgcgc 120

Qy    121 ccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
          |||| ||||||| |||||||||| || ||||||||||||||||||||||||||||||||
Db    121 ccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccat 240
          |||||| |||||||||| || ||||| ||||| | ||||||||||||||||||| || ||
Db    181 aaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcac 240

Qy    241 cgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatc 300
          || ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300

Qy    301 gagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngc 360
          |||||||||||||||  || |||||| ||||||| || ||| ||||  || | |  ||
```

```
Db    301 gagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgc 360

Qy    361 aatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccct 420
          || ||||| ||||| || || |||||||||||| || | ||||||||||||||||||||||
Db    361 aacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccct 420

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctngcg 480
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    421 gccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480

Qy    481 gcagaccctgtggttgtgatcaaagaagccgncgangatttcagacgangtntgctggta 540
          || |||||| ||| | |||||||||||||| ||| || |||| ||| |  || ||||||
Db    481 gccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggta 540

Qy    541 ttgagccgtttgaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgccg 600
          |||||||| ||||||| |||||||||||||||||||| ||||| ||||||||||||||||
Db    541 ttgagccgcctgaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgccg 600

Qy    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatccggctggtttg 660
          |||||||||||||||| ||||||||||||||||||| |||||||||||||| ||||| |||
Db    601 tctgaaaatgctgccaatatcgaaacacatgaatttggcggcccgcatcctgccggcttg 660

Qy    661 agtggcacgcacattcatttcattgagccggtcggtgcaaacaaaaccgtttggaccatc 720
          |||||||||||||||||||||| ||||| ||||| || || |||||||| |||||||||
Db    661 agtggcacgcacattcatttcatcgagccagtcggcgcgaataaaaccgtgtggaccatc 720

Qy    721 aattatcaagatgtaattgccatcggacgtttgtttgcaacaggccgtctgaacaccgag 780
          |||||||||| || ||||| |||||||||||||| | |||||||||||||| ||||||
Db    721 aattatcaagacgtgattgctatcggacgtttgttcgtaacaggccgtctgaataccgag 780

Qy    781 cgcgtgattgctttgggtggtttctcaagtcaacaaaccacgcctcttgcgtaccgttttg 840
          |||||| |||| ||||| ||    |||||||||||||| ||||||||||||||||||||
Db    781 cgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840

Qy    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacgcagacaaccgcgtgatt 900
          |||||||| || || ||| |||| || ||||||||||||||||| ||||||||||||||
Db    841 ggtgcgaaggtgtctcaacttaccgccggcgaattggttgacgcggacaaccgcgtgatt 900

Qy    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960
          |||||||||||||||||||| |||||| |||||||||||| ||||||||||||||||||||
Db    901 tccggttcggtattgaacggtgcgattgcacaaggcgcgcatgattatttgggacgctac 960

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020

Qy   1021 ccgcagccggacaaatactccatcacgcgtacgaccctcggccatttcctgaaaaacaaa 1080
          |||||||||||||||||||||||||||||| || || |||||||||||||| ||||||||
Db   1021 ccgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaa 1080

Qy   1081 ctcttcaagttcacgacagccgtcaacggtggcgaccgcgccatggtgccgattggtact 1140
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||| |||||
Db   1081 ctcttcaagttcacgacagccgtcaacggcggcgaccgcgccatggtaccgatcggcact 1140

Qy   1141 tacgagcgcgtaatgccgctagacatcctgcctaccctgcttttgcgcgatttaatcgtc 1200
          || |||||||||||||| | |||||||||||||| |||||||||||||||||||||||||
Db   1141 tatgagcgcgtaatgccgttggacatcctgcctaccttgcttttgcgcgatttaatcgtc 1200

Qy   1201 ggcgataccgacagcgcgcaagcattgggttgcttggaattggacgaagaagacctcgct 1260
          |||||||||||||||||| || ||||||||||||||||||||||||||||||||||||||
Db   1201 ggcgataccgacagcgcgcaggctttgggttgcttggaattggacgaagaagacctcgct 1260

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatanggcccgctgttgcgtaaggtgctg 1320
          |||||||||||||||||||||||||||||||||||||| |||||||||||||| || ||||||
Db   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320

Qy   1321 gaaaccnttgagaaggaaggctga 1344
          |||||| ||||||||||||||||||
Db   1321 gaaaccattgagaaggaaggctga 1344

RESULT   6
AAT42063
```

```
ID   AAT42063 standard; DNA; 1830121 BP.
XX
AC   AAT42063;
XX
DT   14-SEP-1999  (first entry)
XX
DE   Haemophilus influenzae complete genome sequence.
XX
KW   Genome; bacterium; Haemophilus influenzae; computer readable medium;
KW   expression modulating fragment; regulation; gene expression; vector;
KW   organism; open reading frame; ORF; ds.
XX
OS   Haemophilus influenzae.
XX
PN   WO9633276-A1.
XX
PD   24-OCT-1996.
XX
PF   22-APR-1996;   96WO-US05320.
XX
PR   07-JUN-1995;   95US-0487429.
PR   21-APR-1995;   95US-0426787.
PR   07-JUN-1995;   95US-0476102.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Adams MD,  Fleischmann RD,  Smith HO,  Venter JC,  White O;
XX
DR   WPI; 1996-485782/48.
XX
PT   Haemophilus influenzae Rd genome recorded on computer readable
PT   medium  - useful for identifying commercially important nucleic acid
PT   fragments by homology searching
XX
PS   Claim 1; Page 77.2-77.1091; 1291pp; English.
XX
CC   This sequence represents the complete genome sequence of the bacterium
CC   Haemophilus influenzae strain Rd.  The invention relates to a computer
CC   readable medium (CRM) having recorded upon it the complete H.influenzae
CC   nucleotide sequence (I), a representative fragment of (I) or a nucleotide
CC   sequence at least 99% identical to (I).  By providing the full-length
CC   genomic sequence in a computer readable form, it is possible to identify
CC   commercially important nucleic acid fragments and expression modulating
CC   fragments (EMFs) of the Haemophilus genome. The EMFs can be used to
CC   regulate the expression of a nucleic acid molecule.  Vectors and altered
CC   organisms comprising the predicted ORFS can be used to produce any of the
CC   polypeptide fragments of the H. influenzae Rd genome.
XX
SQ   Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;
```

```
  Query Match             39.2%;  Score 527.2;  DB 17;  Length 1830121;
  Best Local Similarity   62.5%;  Pred. No. 6e-150;
  Matches  842;  Conservative    0;  Mismatches  500;  Indels    5;  Gaps    2;

Qy       1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcatt 60
           ||||||| ||| || |||||| |  | || || || ||||| | ||| |  ||||| ||
Db  179007 atgattacaattaagaaaggtttggatctcccaattgcgggaaaaccagcacaagtaatc 179066

Qy      61 tatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
            ||  |||   ||   | |   |||| ||| | || || || || ||||  || |||||
Db  179067 catagcggcaacgctgtgaatcaagttgcgattctaggtgaggagtatgtggggatgcgt 179126

Qy     121 ccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
           || |  ||||| ||    |||||||||  || |||||||| ||||| || |||||||||
Db  179127 ccttcaatgaaggtgcgcgaaggcgatgttgtgaaaaaaggtcaagtacttttttgaagac 179186

Qy     181 aaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccat 240
           |||||  |||| || ||  | ||||| || || |     || |  ||| | || ||| ||
Db  179187 aaaaaaaatcctggtgtaatttttacagcccctgcaagcggtaccatcactgcaatcaat 179246

Qy     241 cgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatc 300
```

```
Db 179247 cgtggcgaaaaacgtgtattacaatctgttgttattaatgtagaaggtgatgaaaaaatc 179306

Qy    301 gagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngc 360
Db 179307 acttttgcaaaatatagcacagaacaattgaatacactttcttctgaacaagtaaaacaa 179366

Qy    361 aatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccct 420
Db 179367 aacttgatagaatctggcttatggacggcattacgtactcgtccatttagtaaagttcca 179426

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctngcg 480
Db 179427 tccattgaaagtgaagcatcttctatttttgtaaatgcgatggataccaatccattagca 179486

Qy    481 gcagaccctgtggttgtgatcaaagaagccgncgangatttcagacgangtntgctggta 540
Db 179487 gccgatccctctgtcgtattaaaagaatattcgcaagacttcactaacggtttaaccgta 179546

Qy    541 ttgagccgtttgaccgagcgtaa---aatccatgtgtgtaaggcagctggcgcagacgtg 597
Db 179547 ttaagtcgtctattcccttcaaaaccattacacttatgtaaagcaggcgattctaatatt 179606

Qy    598 ccgtctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatccggccggt 657
Db 179607 ccaactgctgatcttgaaaatttacagattcatgattttactggtgttcatccagctggt 179666

Qy    658 ttgagtggcacgcacattcatttcattgagccggtcggtgcaaacaaaaccgtttggacc 717
Db 179667 cttgtagggacgcatattcactttattgatcctgtcggtattcaaaaaactgtatggcat 179726

Qy    718 atcaattatcaagatgtaattgccatcggacgtttgtttgcaacaggccgtctgaacacc 777
Db 179727 atcaattatcaagatgtgatcgctgtaggtaaattatttacaacaggcgaactttattca 179786

Qy    778 gagcgcgtgattgctttgggtggttctcaagtcaacaaaccacgcctcttgcgtaccgtt 837
Db 179787 gaacgagtaatttctcttgcaggtcctcaagtgaaagaacctcgtttagtacgtacaaca 179846

Qy    838 ttgggtgcgaaagtatcgcaaattactgcgggcgaattggttgacgcagacaaccgcgtg 897
Db 179847 atcggtgctaatctttctcaattaacccaaaatgaattaagtgcaggtaaaaatcgtgta 179906

Qy    898 atttccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgc 957
Db 179907 atttctggctcagttctttgtggtcaaatagcaaaagattcacatgattatttaggtcgc 179966

Qy    958 taccacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggtt 1017
Db 179967 tatgctttgcaagtatcagtaatcgctgaaggtaatgaaaaagaatttttggttggatt 180026

Qy   1018 gcgccgcagccggacaaatactccatcacgcgtacgaccctcggccatttcctgaaaaac 1077
Db 180027 atgccacaagccaataaatattctgtgactcgtaccgtattagggccacttcagtaaa-- 180084

Qy   1078 aaactcttcaagttcacgacagccgtcaacggtggcgaccgcgccatggtgccgattggt 1137
Db 180085 aaattgtttaatttactacttcagaaaatggtggcgaacgtgcaatggtgccaatcggt 180144

Qy   1138 acttacgagcgcgtaatgccgctagacatcctgcctaccctgcttttgcgcgatttaatc 1197
Db 180145 agctatgagcgcgtaatgccattggacattttacctacattattattacgtgatttaatc 180204

Qy   1198 gtcggcgataccgacagcgcgcaagcattgggttgcttggaattggacgaagaagacctc 1257
Db 180205 gtgggcgataccgatggtgctcaagagttaggttgtctggaattagacgaagaagacttg 180264

Qy   1258 gctttgtgcagcttcgtctgcccgggcaaatacgaatanggcccgctgttgcgtaaggtg 1317
Db 180265 gcgttatgttctttcgtttgtccaggcaaatatgaatacggttcaatcttgcgtcaagtc 180324

Qy   1318 ctggaaaccnttgagaaggaaggctga 1344
```

Db 180325 ttagataagattgagaaggaaggttaa 180351

RESULT 7
AAZ12025
ID AAZ12025 standard; DNA; 474 BP.
XX
AC AAZ12025;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria meningitidis partial ORF22 sequence.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS Neisseria meningitidis.
XX
PN WO9924578-A2.
XX
PD 20-MAY-1999.
XX
PF 09-OCT-1998; 98WO-IB01665.
XX
PR 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX
DR WPI; 1999-327407/27.
DR P-PSDB; AAY38560.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
PS Claim 9; Page 122; 524pp; English.
XX
CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicaemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX
SQ Sequence 474 BP; 122 A; 130 C; 130 G; 91 T; 1 other;

Query Match 32.1%; Score 431.6; DB 20; Length 474;
Best Local Similarity 93.5%; Pred. No. 2.4e-122;
Matches 443; Conservative 1; Mismatches 30; Indels 0; Gaps 0;

```
Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcatt 60
Db      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60

Qy     61 tatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
Db     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Qy    121 ccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
Db    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccat 240
```

```
Db    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240

Qy    241 cgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatc 300
Db    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaargcaacgacgaaatc 300

Qy    301 gagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngc 360
Db    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360

Qy    361 aatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccct 420
Db    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
Db    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
```

RESULT 8
AAA81335
ID AAA81335 standard; DNA; 474 BP.
XX
AC AAA81335;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis MenB polynucleotide sequence ORF number 14.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Disclosure; Page 203; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neissariae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
SQ Sequence 474 BP; 122 A; 130 C; 130 G; 91 T; 1 other;

Query Match 32.1%; Score 431.6; DB 21; Length 474;
Best Local Similarity 93.5%; Pred. No. 2.4e-122;
Matches 443; Conservative 1; Mismatches 30; Indels 0; Gaps 0;

```
Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcatt 60
Db      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60

Qy     61 tatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
Db     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Qy    121 ccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
Db    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccat 240
Db    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240

Qy    241 cgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatc 300
Db    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaargcaacgacgaaatc 300

Qy    301 gagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngc 360
Db    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360

Qy    361 aatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccct 420
Db    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
Db    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
```

RESULT 9
AAZ54035/c
ID AAZ54035 standard; DNA; 363 BP.
XX
AC AAZ54035;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 628 partial DNA sequence SEQ ID NO:2019.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.

```
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY75273.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics -
XX
PS   Claim 7; Page 1003-1004; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 363 BP; 75 A; 105 C; 103 G; 80 T; 0 other;

  Query Match             26.1%;  Score 351;  DB 21;  Length 363;
  Best Local Similarity   96.7%;  Pred. No. 1.3e-97;
  Matches  351;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    312 ctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngcaatctgatcca 371
          ||||||||||||||||||||||||||||||||||| ||| |  |  ||||||||||||||
Db    363 CTACGCGCCCGAAGCGTTGGCAAACTTAAGCGGCGAAGAAGTGCGCCGCAATCTGATCCA 304

Qy    372 atccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccctgccgtcgatgc 431
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    303 ATCCGGTTTGTGGACTGCGCTGCGTACCCGTCCGTTCAGCAAAATCCCTGCCGTCGATGC 244

Qy    432 cgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctngcggcagaccctgt 491
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    243 CGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCGGCAGACCCTGT 184

Qy    492 ggttgtgatcaaagaagccgncgangatttcagacgangtntgctggtattgagccgttt 551
          |||||||||||||||||||| ||| |||||||||||| || ||||||||||||||||||||
Db    183 GGTTGTGATCAAAGAAGCCGCCGAGGATTTCAGACGAGGTTTGCTGGTATTGAGCCGTTT 124

Qy    552 gaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgccgtctgaaaatgc 611
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 GACCGAGCGTAAAATCCATGTGTGTAAGGCAGCTGGCGCAGACGTGCCGTCTGAAAATGC 64

Qy    612 tgccaacatcgaaacacatgaattcggcggcccgcatccggccggtttgagtggcacgca 671
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 TGCCAACATCGAAACACATGAATTCGGCGGCCCGCATCCGGCCGGTTTGAGTGGCACGCA 4

Qy    672 cat 674
          |||
Db      3 CAT 1
```

```
RESULT  10
AAF28548/c
ID   AAF28548 standard; DNA; 96109 BP.
XX
AC   AAF28548;
XX
DT   04-APR-2001  (first entry)
XX
DE   Genomic fragment #35.
XX
KW   Genomic library; bacteria; human upper airway; otitis media; sinusitis;
KW   bronchopulmonary; endocarditis; meningitis; ss.
XX
OS   Moraxella catarrhalis.
XX
PN   WO200078968-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US16649.
XX
PR   18-JUN-1999;   99US-0140121.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lagace RE,  Patterson C,  Berg KL;
XX
DR   WPI; 2001-041427/05.
XX
PT   Genomic library for identifying diagnostic and therapeutic
PT   compositions, and for identifying virulence factors, regulatory
PT   elements and drug targets, comprises Moraxella catarrhalis nucleic
PT   acids -
XX
PS   Claim 1; Page 345-368; 545pp; English.
XX
CC   The present invention relates to a Moraxella catarrhalis genomic library
CC   comprising of a combination of 41 nucleic acid molecules (see
CC   AAF28514-AAF28554). The library has a number of uses described in the
CC   specification e.g. is useful for aidentifying diagnostic and therapeutic
CC   compositions. M. catarrhalis (Branhamella catarrhalis) is a large
CC   aerobic, gram-negative diplococcus, normally found among the bacterial
CC   flora of human upper airways. M. catarrhalis is known to cause acute,
CC   localised infections such as otitis media, sinusitis and bronchopulmonary
CC   infection and life-threatening, systemic diseases including endocarditis
CC   and meningitis.
XX
SQ   Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 other;

  Query Match             24.8%;  Score 333.6;  DB 22;  Length 96109;
  Best Local Similarity   54.6%;  Pred. No. 4.6e-91;
  Matches  742;  Conservative    0;  Mismatches  596;  Indels   21;  Gaps    4;

Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcatt 60
          |||||||  ||||||||||||  |  | ||||| ||| | |||   ||    |    ||
Db  85201 ATGATTACCATCAAAAAAGGCTTGGATCTGCCAATCACTGGCGAGCCCTCCAAGCAGAT- 85141

Qy     61 tatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
            |  || || |       ||  | || |||||  | ||| | || |||| ||| |||||
Db  85142 --TAGCGAGCACACGCCCACTAAGGTGGCGTTAGTCGGCTACGATTATGTCGGCATGCGT 85085

Qy    121 ccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
          ||    ||| | ||||| |||| |||   ||||  ||||| || ||||| |  ||||||
Db  85084 CCGACCATGCATGTCAAAGAAGGTGATAAGGTCACCAAAGGACAGGTGCTTTACGAAGAC 85025

Qy    181 aaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccat 240
          |||||     | |||||  | | ||| || || |||   ||||   |||  ||  || ||
Db  85024 AAAAAACGAGCAGGCGTCCTTTATACAGCACCAGTTGGCGGCACTGTCGTTGCTGTCAAT 84965

Qy    241 cgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttg------aaggcaacgac 294
```

```
Db  84964 CGTGGCGAACGACGAGTATTTGAAAGCTTGGTCATCGCCGTCGATCCAAATGCAAGTGAA 84905

Qy    295 gaaatcgagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantn 354
Db  84904 GAAATCACTTTTAATGCTTATCGCTCTGAAGATTTGTCAATGCTGGATCGCAAGACAGTT 84845

Qy    355 ngnngcaatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaa 414
Db  84844 GTTGAGCAGTTGGTCGCTTCAGGTGAATGGACCGCCCTACGCACTCGTCCTTTTAGCCGT 84785

Qy    415 atccctgccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
Db  84784 ACACCAGAAATTAATAGCATACCTTCTGCCATTTTTGTGACGGCAACAGATACCAATCCT 84725

Qy    475 ctngcggcagaccctgtggttgtgatcaaagaagccgncgangatttcagacgangtntg 534
Db  84724 TTGGCTGTTGACCCTGCCACCGTATTATTTGAAGAAATACAAGCATTTAATGATGGCTTG 84665

Qy    535 ctggtattgagccgtttgaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagac 594
Db  84664 GCAATTTTATCAACGCTATCC---CCAAAAACCTATGTCTGTCATGGTGATATTGTACCT 84608

Qy    595 gtgccgtctgaaaatgctgccaac---------atcgaaacacatgaattcggcggcccg 645
Db  84607 GAGACCTTAAGTAAAACCGCTGACGGTAACGATACCATCTATGAAGGGTTTACTGGTGTA 84548

Qy    646 catccggccggtttgagtggcacgcacattcatttcattgagccggtcggtgcaaacaaa 705
Db  84547 CATCCAGCAGGCAACGCAGGTACGCACATTCACTTTTTACACCCTTTGGCACGAGGCAGC 84488

Qy    706 accgtttggaccatcaattatcaagatgtaattgccatcggacgtttgtttgcaacaggc 765
Db  84487 TTAGTTTGGACGATTAATTATCAAGATGTCATCGCCATTGGTAAGCTATTTACCACAGGT 84428

Qy    766 cgtctgaacaccgagcgcgtgattgctttgggtggttctcaagtcaacaaaccacgcctc 825
Db  84427 CGCATTTACACAGAGCGTCTGATCAGTTTGGCAGGTCCAGCAGTGGCCAATCCTCGCTTG 84368

Qy    826 ttgcgtaccgttttgggtgcgaaagtatcgcaaattactgcgggcgaattggttgacgca 885
Db  84367 GTGCGTACCACTCGAGGTGCTGATTTGACCGAATTGACCCAAGGTGAACTCAACGGCGAC 84308

Qy    886 gacaaccgcgtgatttccggttcggtattgaacggcgcgattacacaaggcgcgcacgat 945
Db  84307 GATAACCGTGTCATTTCAGGCTCTGTTCTATCAGGTCGCACAGCTGTTGGTGCTACGGCA 84248

Qy    946 tatttgggacgctaccacaatcagatttccgttatcgaagaaggccgcagcaaagagctg 1005
Db  84247 TTTTTGGGCCGCTTTCATAATCAAGTCTCTGTACTTACCGAAGGTCGTGAACGCCCTGCA 84188

Qy   1006 ttcggctgggttgcgccgcagccggacaaatactccatcacgcgtacgaccctcggccat 1065
Db  84187 CTTCACTTTTTTACACCAGGGGCAAATCGTTTTTCAAAATTGCCAATTTATATTTCACAG 84128

Qy   1066 ttcctgaaaaacaaactcttcaagttcacgacagccgtcaacggtggcgaccgcgccatg 1125
Db  84127 TTTTTTAAAGGCAAAAAATACGATTTTACGACCAGTACCAACGGCTCGCCTCGTGCCATG 84068

Qy   1126 gtgccgattggtacttacgagcgcgtaatgccgctagacatcctgcctaccctgcttttg 1185
Db  84067 GTACCAATCGGCTTATTTGAAACTGTCATGCCTCAAGACTATCTACCTACCCAGCTACTG 84008

Qy   1186 cgcgatttaatcgtcggcgataccgacagcgcgcaagcattgggttgcttggaattggac 1245
Db  84007 CGTGCTTTGATTGTCGAAGATATCATCAGTGCGGTTGAATTGGGTGCACTTGAGCTTGAT 83948

Qy   1246 gaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatanggcccgctg 1305
Db  83947 GAAGAAGATTTGGCACTGTGTACTTTCGTATCGCCAGGCAAATATGAGTTCGGTGATATT 83888

Qy   1306 ttgcgtaaggtgctggaaaccnttgagaaggaaggctga 1344
```

```
Db  83887 TTGCGTGATAACTTGACCCGTATTGAGCAGGAAGGCTAA 83849


RESULT  11
AAZ54034/c
ID   AAZ54034 standard; DNA; 363 BP.
XX
AC   AAZ54034;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 628 partial DNA sequence SEQ ID NO:2017.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY75272.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics -
XX
PS   Claim 7; Page 1003; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 363 BP; 74 A; 96 C; 110 G; 83 T; 0 other;


  Query Match             23.6%;  Score 317.4;  DB 21;  Length 363;
  Best Local Similarity   90.9%;  Pred. No. 2.8e-87;
  Matches  330;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy    312 ctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngcaatctgatcca 371
Db    363 CTACGCACCTGAAGCGCTGGCAAACTTAAGCGGCGAAGAAGTGCGCCGCAACCTGATCCA 304
```

```
Qy    372 atccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccctgccgtcgatgc 431
          ||||||||||||||||||||||||| | |||||||||||||||||| ||||||||||||||
Db    303 ATCCGGTTTGTGGACTGCGCTGCGCACCCGTCCGTTCAGCAAAATTCCTGCCGTCGATGC 244

Qy    432 cgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctngcggcagaccctgt 491
          ||||||||||||||||||||||||||||||||||||||||||||||| || || ||||||
Db    243 CGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCTGCCGACCCTAC 184

Qy    492 ggttgtgatcaaagaagccgncgangatttcagacgangtntgctggtattgagccgttt 551
          |||  | |||||||||||| ||| ||||||| |||  |  || ||||||||||||||||
Db    183 GGTCATTATCAAAGAAGCCGCCGAGGATTTCAAACGCGGCCTGTTGGTATTGAGCCGTTT 124

Qy    552 gaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgccgtctgaaaatgc 611
          ||||||  || |||||||||| ||||||||||||||||||||||||||||||||||||||
Db    123 GACCGAACGCAAAATCCATGTTTGTAAGGCAGCTGGCGCAGACGTGCCGTCTGAAAATGC 64

Qy    612 tgccaacatcgaaacacatgaattcggcggcccgcatccggccggtttgagtggcacgca 671
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     63 TGCCAACATCGAAACACATGAATTCGGCGGCCCGCATCCTGCCGGTTTGAGTGGCACGCA 4

Qy    672 cat 674
          |||
Db      3 CAT 1
```

RESULT 12
AAZ54033/c
ID AAZ54033 standard; DNA; 363 BP.
XX
AC AAZ54033;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria gonorrheae ORF 628 partial DNA sequence SEQ ID NO:2015.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria gonorrheae.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
DR P-PSDB; AAY75271.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 7; Page 1003; 1453pp; English.
XX
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 363 BP; 71 A; 92 C; 109 G; 91 T; 0 other;

```
  Query Match             21.7%;  Score 291.8;  DB 21;  Length 363;
  Best Local Similarity   86.5%;  Pred. No. 2.1e-79;
  Matches  314;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy    312 ctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngcaatctgatcca 371
          |||||  || |||||| ||||||| || ||| ||||  || |  |  |||| ||||| ||
Db    363 CTACGTACCTGAAGCGCTGGCAAAATTGAGCAGCGAAAAAGTGCGCCGCAACCTGATTCA 304

Qy    372 atccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccctgccgtcgatgc 431
          ||| || || |||||||||||| || | ||||||||||||||||||||||||||||||||
Db    303 ATCAGGCTTATGGACTGCGCTTCGCACCCGTCCGTTCAGCAAAATCCCTGCCGTAGATGC 244

Qy    432 cgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctngcggcagaccctgt 491
          ||||||||||||||||||||||||||||||||||||||||||||||| || || ||||||
Db    243 CGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCTGCCGACCCTAC 184

Qy    492 ggttgtgatcaaagaagccgncgangatttcagacgangtntgctggtattgagccgttt 551
          |||  | |||||||||||| ||| || |||| |||  |  || ||||||||||||  |
Db    183 GGTCATCATCAAAGAAGCCGCCGAAGACTTCAAACGCGGCCTGTTGGTATTGAGCCGCCT 124

Qy    552 gaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgccgtctgaaaatgc 611
          ||||||  |||||||||||||||||||| ||||| ||||||||||||||||||||||||
Db    123 GACCGAACGTAAAATCCATGTGTGTAAAGCAGCAGGCGCAGACGTGCCGTCTGAAAATGC 64

Qy    612 tgccaacatcgaaacacatgaattcggcggcccgcatccggccggtttgagtggcacgca 671
          ||||||  |||||||||||||||||  |||||||||||||  ||||| ||||||||||||
Db     63 TGCCAATATCGAAACACATGAATTTGGCGGCCCGCATCCTGCCGGCTTGAGTGGCACGCA 4

Qy    672 cat 674
          |||
Db      3 CAT 1
```

RESULT 13
AAX91657
ID AAX91657 standard; DNA; 1353 BP.
XX
AC AAX91657;
XX
DT 25-AUG-1999 (first entry)
XX
DE Porphorymonas gingivalis protein PG1 encoding DNA.
XX
KW Porphorymonas gingivalis; PG; periodontal disease; gingivitis;
KW vaccine; antigenic; ds.
XX
OS Porphorymonas gingivalis.
XX
PN WO9929870-A1.
XX
PD 17-JUN-1999.
XX
PF 10-DEC-1998; 98WO-AU01023.
XX
PR 04-AUG-1998; 98AU-0005028.
PR 10-DEC-1997; 97AU-0000839.
PR 31-DEC-1997; 97AU-0001182.

```
PR   30-JAN-1998;   98AU-0001546.
PR   10-MAR-1998;   98AU-0002264.
PR   09-APR-1998;   98AU-0002911.
PR   23-APR-1998;   98AU-0003128.
PR   05-MAY-1998;   98AU-0003338.
PR   22-MAY-1998;   98AU-0003654.
PR   29-JUL-1998;   98AU-0004917.
XX
PA   (CSLC-) CSL LTD.
XX
PI   Agius CT,  Barr IG,  Hocking DM,  Margetts MB,  Patterson MA;
PI   Ross BC,  Rothel LJ,  Webb EA;
XX
DR   WPI; 1999-385613/32.
DR   P-PSDB; AAY34439.
XX
PT   Antigenic Porphorymonas gingivalis peptides for preventing
PT   gingivitis
XX
PS   Claim 12; Page 179; 588pp; English.
XX
CC   AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC   Porphorymonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC   AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC   isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC   activity with a vaccine mechanism of action. The PG polypeptides can be
CC   used as vaccines especially against Porphorymonas gingivalis. Probes can
CC   be used to detect Porphorymonas gingivalis in standard hybridisation
CC   assays. Porphorymonas gingivalis is involved in periodontal disease
CC   especially gingivitis.
XX
SQ   Sequence 1353 BP; 337 A; 333 C; 379 G; 304 T; 0 other;
```

```
  Query Match             9.1%;  Score 122.2;  DB 20;  Length 1353;
  Best Local Similarity   46.0%;  Pred. No. 5.5e-27;
  Matches  622;  Conservative    0;  Mismatches  700;  Indels   30;  Gaps    6;

Qy       2 tgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcattt 61
           | || ||||  |||||||| ||    ||     |    || | ||||  ||  |  ||
Db      11 ttataaaaacaaaaaaaggccttgcacttaatctgaaaggaaaaccgctgcccgagatgc 70

Qy      62 atgacgggcccgtc---attaccgaagtcgcgttgcttggcgaagaatatgccggtatgc 118
             | ||  || | |   | | |      |||| |  |   ||| || | ||  ||| |
Db      71 tggccgaaccggcccaaagtcctacttacgcggtcgtgcccgacgattttgaaggtgtta 130

Qy     119 gcccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaag 178
           ||||  | |||  |         || |||   ||       |||  ||  ||| |  |
Db     131 tccccaaggtgacggctcgtccgggggataaggtgcgtgccggctcagcactgatgcacc 190

Qy     179 acaaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatcc 238
           ||||     ||||||   ||  ||||||    || |||   ||| || |   |||  |
Db     191 acaaggcatatccggagatgaagtttacaagtccggttagcggcgaagtgatcgcggtga 250

Qy     239 atcgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaa 298
           ||||||| |  ||||||       ||  | |    || |  ||  ||    |||     |
Db     251 atcgcggtgccaagcgc------aaggtgttgagcatcgaggtgaaaccggacggactga 304

Qy     299 tcgagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnn 358
           ||| | |||    | | |   ||  |    | |   | |    | |||  |  |
Db     305 acgaatacgagtcattccctgtcggggatccgtctgccctctctgccgaacagatcaagg 364

Qy     359 gcaatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatcc 418
             | |  |      |||| |||||  |    |       |||||  | |  || | |
Db     365 agcttttactgtcgagcggtatgtggggttttattaagcaacgtccttacgacatagtgg 424

Qy     419 ctgccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctng 478
           || |    |||   |  ||   || ||| |   | | ||| |     ||    ||  | |
Db     425 ctacaccggatatagctccacgcgacatttatattactgccaactttactgcaccattgg 484

Qy     479 cggcagaccctgtggttgtgatcaaagaagccgncgangatttcagacgangtntgctgg 538
           |  | |||   |   |  |  |   || ||  |     |   |          |    |
Db     485 ctccggacttcgatttcatcgttcgaggagaagaacgcgccctgcagactgccatcgatg 544

Qy     539 tattgagccgtttgaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgc 598
             |||  |    | ||     | ||  |  ||||| |   |    | ||  ||    ||
Db     545 ccttggccaaactcacgacaggaaaggtgtatgtgggcctgaagccgggttcatctctgg 604

Qy     599 cgtctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatccggccggtt 658
             |   | |||||  |  | | |  |||  |||          ||| || ||||||||  |||
Db     605 gcttgcacaatgcagaaatcgtagaagtaca---------cggacctcatccggcaggta 655

Qy     659 tgagtggcacgcacattcatttcattgagccggtcggtgcaaacaaaaccgtttggacca 718
             |||  ||  ||  ||   |   ||||  ||  |     | |||| || |||||
Db     656 acgtgggcgtgctgatcaatcatacgaagccaatcaatcggggcgaaacggtgtggacgc 715

Qy     719 tcaattatcaagatgtaattgccatcggacgtttgtttgcaacaggccgtctgaacaccg 778
           ||||   |   ||  | || |  ||||||||||||  |    || |||       |
Db     716 tcaaggctaccgacctgatcgtgatcggacgtttcctgcttacgggcaaagccgatttta 775

Qy     779 agcgcgtgattgctttgggtggttctcaagtcaacaaaccacgcctcttgcgtaccgttt 838
             | |||||||  ||   || ||  | |         |    |   | | ||||   |
Db     776 ccagaatgattgccatgaccggctcagacgctgcagctcacggatacgtccgtattatgc 835

Qy     839 tgggtgcgaaagtatcgcaaattactgcgggcgaattg------gttgacgcagacaacc 892
           ||||  || || |           | |||  | ||        ||  |  || | |
Db     836 cgggttgcaatgtctttgcttccttccccggccgactgacaataaaggaatctcacgagc 895

Qy     893 gcgtgatttccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgg 952
           | |||||    ||    ||  ||| |||   ||         |    | |   ||   ||
Db     896 gtgtgatcgatggcaatgtgctgaccggtaagaagctctgcgagaaggagcctttcctgt 955

Qy     953 gacgctaccacaatcagatttccgttatcgaagaaggccgc---agcaaagagctgttcg 1009
             | |       | |||||  | || |||   ||||||  |       | || || ||||
Db     956 cagcccggtgtgaccagatcacggtgatccccgaaggcgacgatgtggacgaactcttcg 1015

Qy    1010 gctgggttgcgccgcagccggacaaatactccatcacgcgtacgaccctcggccatttcc 1069
           | |||| ||| || |  |  ||  | |||  ||| |   |   |  |     ||      || |
Db    1016 ggtgggctgcaccccgtctcgatcagtacagcatgagcagagcttatttctcttggttgc 1075

Qy    1070 ---tgaaaaacaaactcttcaagttcacgacagccgtcaacggtggcgaccgcgccatgg 1126
              ||||||||||   | |    ||    |     |||| |||||||| || || |||
Db    1076 aggggaaaaacaaagagtacgtactcgatgcccggatcaagggtggcgaacgtgctatga 1135

Qy    1127 tgccgattggtacttacgagcgcgtaatgccgctagacatcctgcctaccctgcttttgc 1186
           |   ||         || || |||||  | ||| | ||||||   ||        |  |
Db    1136 tcatgagcaacgagtatgaccgcgttttcccgatggacatctatccggagtatttgctca 1195

Qy    1187 gcgatttaatcgtcggcgataccgacagcgcgcaagcattgggttgcttggaattggacg 1246
             | | | || |    ||| | |||||    | | |  || ||     |  ||| |||
Db    1196 aggctattatagcattcgacatcgacaagatggaggacttaggcatatatgaagtggctc 1255

Qy    1247 aagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatanggcccgctgt 1306
             || ||| | ||    |||   || ||      |   |||   |||       |   |
Db    1256 cggaggactttgccacttgcgaatttgtggatacatccaagatcgagctgcagcgtatcg 1315

Qy    1307 tgcgtaaggtgctggaaaccnttgagaaggaa 1338
           | ||  |||   |||| |   |  | ||||||
Db    1316 ttcgcgagggcttggatatgctctataaggaa 1347
```

```
RESULT  14
AAX91536
ID   AAX91536 standard; DNA; 1362 BP.
XX
AC   AAX91536;
XX
DT   25-AUG-1999  (first entry)
XX
DE   Porphorymonas gingivalis protein PG1 ORF encoding DNA.
XX
KW   Porphorymonas gingivalis; PG; periodontal disease; gingivitis;
KW   vaccine; antigenic; ds.
XX
OS   Porphorymonas gingivalis.
XX
```

PN WO9929870-A1.
XX
PD 17-JUN-1999.
XX
PF 10-DEC-1998; 98WO-AU01023.
XX
PR 04-AUG-1998; 98AU-0005028.
PR 10-DEC-1997; 97AU-0000839.
PR 31-DEC-1997; 97AU-0001182.
PR 30-JAN-1998; 98AU-0001546.
PR 10-MAR-1998; 98AU-0002264.
PR 09-APR-1998; 98AU-0002911.
PR 23-APR-1998; 98AU-0003128.
PR 05-MAY-1998; 98AU-0003338.
PR 22-MAY-1998; 98AU-0003654.
PR 29-JUL-1998; 98AU-0004917.
XX
PA (CSLC-) CSL LTD.
XX
PI Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;
PI Ross BC, Rothel LJ, Webb EA;
XX
DR WPI; 1999-385613/32.
DR P-PSDB; AAY34318.
XX
PT Antigenic Porphorymonas gingivalis peptides for preventing
PT gingivitis
XX
PS Claim 12; Page 96; 588pp; English.
XX
CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC Porphorymonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC activity with a vaccine mechanism of action. The PG polypeptides can be
CC used as vaccines especially against Porphorymonas gingivalis. Probes can
CC be used to detect Porphorymonas gingivalis in standard hybridisation
CC assays. Porphorymonas gingivalis is involved in periodontal disease
CC especially gingivitis.
XX
SQ Sequence 1362 BP; 337 A; 335 C; 381 G; 309 T; 0 other;

Query Match 9.1%; Score 122.2; DB 20; Length 1362;
Best Local Similarity 46.0%; Pred. No. 5.5e-27;
Matches 622; Conservative 0; Mismatches 700; Indels 30; Gaps 6;

```
Qy       2 tgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcattt 61
Db      20 ttataaaaacaaaaaaaggccttgcacttaatctgaaaggaaaaccgctgcccgagatgc 79

Qy      62 atgacgggcccgtc---attaccgaagtcgcgttgcttggcgaagaatatgccggtatgc 118
Db      80 tggccgaaccggcccaaagtcctacttacgcggtcgtgcccgacgattttgaaggtgtta 139

Qy     119 gcccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaag 178
Db     140 tccccaaggtgacggctcgtccgggggataaggtgcgtgccggctcagcactgatgcacc 199

Qy     179 acaaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatcc 238
Db     200 acaaggcatatccggagatgaagtttacaagtccggttagcggcgaagtgatcgcggtga 259

Qy     239 atcgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaa 298
Db     260 atcgcggtgccaagcgc------aaggtgttgagcatcgaggtgaaaccggacggactga 313

Qy     299 tcgagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnn 358
Db     314 acgaatacgagtcattccctgtcggggatccgtctgccctctctgccgaacagatcaagg 373

Qy     359 gcaatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatcc 418
```

```
Db     374 agcttttactgtcgagcggtatgtggggttttattaagcaacgtccttacgacatagtgg 433

Qy     419 ctgccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctng 478
Db     434 ctacaccggatatagctccacgcgacatttatattactgccaactttactgcaccattgg 493

Qy     479 cggcagaccctgtggttgtgatcaaagaagccgncgangatttcagacgangtntgctgg 538
Db     494 ctccggacttcgatttcatcgttcgaggagaagaacgcgccctgcagactgccatcgatg 553

Qy     539 tattgagccgtttgaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgc 598
Db     554 ccttggccaaactcacgacaggaaaggtgtatgtgggcctgaagccgggttcatctctgg 613

Qy     599 cgtctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatccggccggtt 658
Db     614 gcttgcacaatgcagaaatcgtagaagtaca---------cggacctcatccggcaggta 664

Qy     659 tgagtggcacgcacattcatttcattgagccggtcggtgcaaacaaaaccgtttggacca 718
Db     665 acgtgggcgtgctgatcaatcatacgaagccaatcaatcggggcgaaacggtgtggacgc 724

Qy     719 tcaattatcaagatgtaattgccatcggacgtttgtttgcaacaggccgtctgaacaccg 778
Db     725 tcaaggctaccgacctgatcgtgatcggacgtttcctgcttacgggcaaagccgatttta 784

Qy     779 agcgcgtgattgctttgggtggttctcaagtcaacaaaccacgcctcttgcgtaccgttt 838
Db     785 ccagaatgattgccatgaccggctcagacgctgcagctcacggatacgtccgtattatgc 844

Qy     839 tgggtgcgaaagtatcgcaaattactgcgggcgaattg------gttgacgcagacaacc 892
Db     845 cgggttgcaatgtctttgcttccttccccggccgactgacaataaaggaatctcacgagc 904

Qy     893 gcgtgatttccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgg 952
Db     905 gtgtgatcgatggcaatgtgctgaccggtaagaagctctgcgagaaggagcctttcctgt 964

Qy     953 gacgctaccacaatcagatttccgttatcgaagaaggccgc---agcaaagagctgttcg 1009
Db     965 cagcccggtgtgaccagatcacggtgatccccgaaggcgacgatgtggacgaactcttcg 1024

Qy    1010 gctgggttgcgccgcagccggacaaatactccatcacgcgtacgaccctcggccatttcc 1069
Db    1025 ggtgggctgcaccccgtctcgatcagtacagcatgagcagagcttatttctcttggttgc 1084

Qy    1070 ---tgaaaaacaaactcttcaagttcacgacagccgtcaacggtggcgaccgcgccatgg 1126
Db    1085 aggggaaaaacaaagagtacgtactcgatgcccggatcaagggtggcgaacgtgctatga 1144

Qy    1127 tgccgattggtacttacgagcgcgtaatgccgctagacatcctgcctaccctgcttttgc 1186
Db    1145 tcatgagcaacgagtatgaccgcgttttcccgatggacatctatccggagtatttgctca 1204

Qy    1187 gcgatttaatcgtcggcgataccgacagcgcgcaagcattgggttgcttggaattggacg 1246
Db    1205 aggctattatagcattcgacatcgacaagatggaggacttaggcatatatgaagtggctc 1264

Qy    1247 aagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatanggcccgctgt 1306
Db    1265 cggaggactttgccacttgcgaatttgtggatacatccaagatcgagctgcagcgtatcg 1324

Qy    1307 tgcgtaaggtgctggaaaccnttgagaaggaa 1338
Db    1325 ttcgcgagggcttggatatgctctataaggaa 1356
```

RESULT 15
AAX84332/c
ID AAX84332 standard; DNA; 5059 BP.
XX
AC AAX84332;
XX

```
DT   08-SEP-1999  (first entry)
XX
DE   Stealth virus nucleic acid clone, SEQ ID NO: 24.
XX
KW   Stealth virus; detection; diagnosis; infection; ss.
XX
OS   Stealth virus.
XX
FH   Key             Location/Qualifiers
FT   misc_difference 3605
FT                   /*tag=  a
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3610
FT                   /*tag=  b
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3615
FT                   /*tag=  c
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3630
FT                   /*tag=  d
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3631
FT                   /*tag=  e
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3638
FT                   /*tag=  f
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3641
FT                   /*tag=  g
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3642
FT                   /*tag=  h
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3657
FT                   /*tag=  i
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3659
FT                   /*tag=  j
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3681
FT                   /*tag=  k
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3698
FT                   /*tag=  l
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3703
FT                   /*tag=  m
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3708
FT                   /*tag=  n
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3713
FT                   /*tag=  o
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3720
FT                   /*tag=  p
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3732
FT                   /*tag=  q
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3733
FT                   /*tag=  r
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3740
FT                   /*tag=  s
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3747
FT                   /*tag=  t
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3750
FT                   /*tag=  u
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3751
FT                   /*tag=  v
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3766
FT                   /*tag=  w
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3778
FT                   /*tag=  x
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3792
FT                   /*tag=  y
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3799
FT                   /*tag=  z
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3800
FT                   /*tag=  aa
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3812
FT                   /*tag=  ab
```

```
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3813
FT                   /*tag=  ac
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3818
FT                   /*tag=  ad
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3819
FT                   /*tag=  ae
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3824
FT                   /*tag=  af
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3827
FT                   /*tag=  ag
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3841
FT                   /*tag=  ah
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3853
FT                   /*tag=  ai
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3857
FT                   /*tag=  aj
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3860
FT                   /*tag=  ak
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3863
FT                   /*tag=  al

  Query Match             2.9%;  Score 39.2;  DB 20;  Length 5059;
  Best Local Similarity   16.3%;  Pred. No. 0.35;
  Matches   44;  Conservative  100;  Mismatches  126;  Indels    0;  Gaps    0;

Qy   1004 tgttcggctgggttgcgccgcagccggacaaatactccatcacgcgtacgaccctcggcc 1063
          |::: : :|:: ::::|: | |:: : ::   :  |  :| |  :: |  |    :||
Db    598 TKYHGSYBTRVYDWSSGYAGNASHTSTHSYSTWNATATHTDATASSNAGNANATNYGGNT 539

Qy   1064 atttcctgaaaaacaaactcttcaagttcacgacagccgtcaacggtggcgaccgcgcca 1123
            |:  |      :::: |:  : : |::  : :|: :: :   :  |  : |   : :|
Db    538 GSTBAGTNCNNTKYRVHNTSNMYBHATYHNTSNMARGBSSYSNKKNAGHNSNCANTRAVA 479

Qy   1124 tggtgccgattggtacttacgagcgcgtaatgccgctagacatcctgcctaccctgcttt 1183
          |:| :  : ::    |: ::: : :  :  :|:  ::|:  :| | :   :::  :  |:
Db    478 TKGASDNSSYDANATCYNWHSTSKSNTDGKDGSDCYDASSNHTGCNKKDAWYYANSTRTK 419

Qy   1184 tgcgcgatttaatcgtcggcgataccgacagcgcgcaagcattgggttgcttggaattgg 1243
           :::  |:   |:  :  : |: | || : : :::   |: ::  :::    | | :  |
Db    418 SSYBNTAYGGNAKGNKKNSAGHGANCGGSYBDSYBRGNGHNKKNTKDDAWNAGNANHGNG 359

Qy   1244 acgaagaagacctcgctttgtgcagcttcg 1273
          :  :|:: :|: | :  :|:|:   :|:::
```

```
Db    358 RTYDAKDGRAYATKKASYTSTKTYHYTKSK 329


RESULT  16
AAZ48249
ID   AAZ48249 standard; cDNA; 9762 BP.
XX
AC   AAZ48249;
XX
DT   20-MAR-2000  (first entry)
XX
DE   Rubella virus Cendehill strain cDNA sequence.
XX
KW   Rubella virus; Cendehill strain; p150 protein, p90 protein; C protein;
KW   E1 protein; E2 protein; vaccine; rubella infection; ss.
XX
OS   Rubella virus.
XX
PN   WO9961637-A1.
XX
PD   02-DEC-1999.
XX
PF   21-MAY-1999;   99WO-CA00479.
XX
PR   22-MAY-1998;   98CA-2238660.
XX
PA   (UYBR-) UNIV BRITISH COLUMBIA.
XX
PI   Chantler J,  Lund K;
XX
DR   WPI; 2000-072628/06.
XX
PT   New isolated rubella virus nucleic acids, used to develop products for
PT   use as vaccines against rubella infection -
XX
PS   Claim 5; Page 38-44; 62pp; English.
XX
CC   The invention provides isolated nucleic acid sequences from the Cendehill
CC   strain of rubella virus. The nucleic acid encodes a Cendehill rubella
CC   protein selected from p150, p90, C, E1 and E2. The products can be used
CC   for the production of vaccines against rubella infection. The products
CC   can also be used for the production of large quantities of virus for use
CC   as antigen in enzyme-linked immunosorbent assays to assess human antibody
CC   levels against rubella. The present sequence represents the sequence of
CC   the Rubella virus Cendehill strain cDNA.
XX
SQ   Sequence 9762 BP; 1464 A; 3765 C; 3006 G; 1527 T; 0 other;


  Query Match             2.9%;  Score 38.8;  DB 21;  Length 9762;
  Best Local Similarity   46.1%;  Pred. No. 0.65;
  Matches  124;  Conservative    0;  Mismatches  145;  Indels    0;  Gaps    0;

Qy     55 gtcatttatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggt 114
          | || || || || | ||| |         | |  |  |     ||| ||    |  |
Db   3944 ggcacttgtgccgcgaccgacgaggggctggcccaggcgtactacgacgacctcgaggtg 4003

Qy    115 atgcgcccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgttt 174
            |||| |  |    | | || ||   | |  ||| ||  |  || ||||   |
Db   4004 cgccgcctcggggatgacgccatggcccgggcggccctcgcatcagtccaacgccctcgc 4063

Qy    175 gaagacaaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgcc 234
           ||| |    |  ||    | ||   |   |  ||  | |   | ||||| | | ||  |
Db   4064 aaaggcccttacaatatcagggtatggaacatggccgcaggcgctggcaagaccacccgc 4123

Qy    235 atccatcgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgac 294
          ||||     | |   | |||||  |   |  |  |||       ||  ||    |   ||
Db   4124 atcctcgctgccttcacgcgcgaagacctctacgtctgccccaccaatgcgctcctgcac 4183

Qy    295 gaaatcgagttcgaacgctacgcgcccga 323
          || ||| ||  | ||| |  |||| |||
Db   4184 gagatccaggccaaactccgcgcgcgcga 4212
```

```
RESULT 17
AAD17184
ID   AAD17184 standard; DNA; 65140 BP.
XX
AC   AAD17184;
XX
DT   29-NOV-2001  (first entry)
XX
DE   Streptomyces noursei nys1 DNA of nystatin PKS gene cluster.
XX
KW   Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster;
KW   antifungal; antibiotic; nys1; ds.
XX
OS   Streptomyces noursei.
XX
FH   Key             Location/Qualifiers
FT   CDS             complement (1..1035)
FT                   /*tag= a
FT                   /product= "NysD2 partial protein"
FT                   /note= "CDS does not include stop codon"
FT   CDS             complement (1056..2576)
FT                   /*tag= b
FT                   /product= "NysD1 protein"
FT   CDS             2806..6906
FT                   /*tag= c
FT                   /product= "NysA protein"
FT   CDS             6952..16530
FT                   /*tag= d
FT                   /product= "NysB protein"
FT   CDS             16550..49840
FT                   /*tag= e
FT                   /product= "NysC protein"
FT   CDS             50260..51015
FT                   /*tag= f
FT                   /product= "NysE protein"
FT   CDS             51405..54305
FT                   /*tag= g
FT                   /product= "NysR1 protein"
FT   CDS             54329..57190
FT                   /*tag= h
FT                   /product= "NysR2 protein"
FT                   /note= "CDS does not include start codon"
FT   CDS             57180..59963
FT                   /*tag= i
FT                   /product= "NysR3 protein"
FT   CDS             60415..61047
FT                   /*tag= j
FT                   /product= "NysR4 (short) protein"
FT                   /note= "CDS does not include start codon"
FT   CDS             61736..62497
FT                   /*tag= k
FT                   /product= "NysR5 protein"
FT                   /note= "CDS does not include start codon"
FT   CDS             complement (62551..63615)
FT                   /*tag= l
FT                   /product= "ORF2 protein"
FT                   /note= "CDS does not include start codon"
FT   CDS             63765..64961
FT                   /*tag= m
FT                   /product= "ORF1 protein"
XX
PN   WO200159126-A2.
XX
PD   16-AUG-2001.
XX
PF   08-FEB-2001; 2001WO-GB00509.
XX
PR   08-FEB-2000; 2000GB-0002840.
PR   10-APR-2000; 2000GB-0008786.
PR   14-APR-2000; 2000GB-0009387.
XX
PA   (UYNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.
PA   (SNTF ) SINTEF STIFTELSEN IND TEK FORSK.
PA   (ALPH-) ALPHARMA AS.
PA   (SINV-) SINVENT AS.
PA   (DZIE/) DZIEGLEWSKA H.
PA   (ZOTC/) ZOTCHEV S B.
PA   (SEKU/) SEKUROVA O N.
PA   (FJAE/) FJAERVIK E.
PA   (BRAU/) BRAUTASET T.
PA   (STRO/) STROM A R.
XX
PI   Zotchev SB,  Sekurova ON,  Fjaervik E,  Brautaset T,  Strom AR;
PI   Valla S,  Ellingsen TE,  Sletta H,  Gulliksen O;
XX
DR   WPI; 2001-557614/62.
DR   P-PSDB; AAE10125, AAE10126, AAE10127, AAE10128, AAE10129, AAE10130,
DR   AAE10131, AAE10132, AAE10133, AAE10134, AAE10135, AAE10136, AAE10137.
XX
PT   New nystatin polyketide synthase polynucleotides and polypeptides,
PT   useful as antibiotics and antifungals -
XX
PS   Claim 2; Page 116-151; 266pp; English.
XX
CC   The present invention relates to the cloning and sequencing of the gene
CC   cluster encoding a modular type I polyketide synthase (PKS) enzyme
CC   involved in the biosynthesis of the macrolide antibiotic nystatin.
CC   The nystatin PKS is useful as antifungal antibiotics. The present
CC   sequence is a Streptomyces noursei nys1 DNA of nystatin PKS gene cluster.
XX
SQ   Sequence 65140 BP; 8270 A; 25171 C; 22273 G; 9426 T; 0 other;


  Query Match             2.8%;  Score 38.2;  DB 22;  Length 65140;
  Best Local Similarity   45.6%;  Pred. No. 2.5;
  Matches  118;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy    193 ggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccatcgcggcgaaaag 252
          |||| || | |   ||||   ||   | ||     || | ||  | |   || |    |
Db   5971 ggcgcggagatcttcgcggacgtggaatgcccgcccggcaccgccgacgacgcccccgac 6030

Qy    253 cgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatcgagttcgaacgc 312
          | || ||| ||  | | | || | | ||  |   |  ||| ||  || | | | ||  ||
Db   6031 cacggactgcaccccgccctgctcgacgcggcccggcacgccgccatggcggtggacggc 6090

Qy    313 tacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngcaatctgatccaa 372
            || |||||  || | |     | |  |     | |   |  |   ||    | | |
Db   6091 accgtgcccgtcgcctggcacggcgtccggctgcacgccgtcggcgccaccgcgctgcgg 6150

Qy    373 tccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccctgccgtcgatgcc 432
            | |  |  |  |  |   |     |   | |  |  ||       | |||  | | |
Db   6151 gtccgcatccgccccaccacgaccggcacgctgaccctcaccgcggtcgacgtgcacggc 6210

Qy    433 gagccgttcgccatcttcg 451
          | |||| ||| || | |||
Db   6211 gcgccggtcgtcaccgtcg 6229


RESULT 18
AAD17186
ID   AAD17186 standard; DNA; 125401 BP.
XX
AC   AAD17186;
XX
DT   29-NOV-2001  (first entry)
XX
DE   Streptomyces noursei nystatin PKS gene cluster DNA.
XX
KW   Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster;
KW   antifungal; antibiotic; ds.
XX
OS   Streptomyces noursei.
XX
FH   Key             Location/Qualifiers
FT   CDS             6337..34771
```

```
FT                   /*tag= a
FT                   /product= "NysI complete protein"
FT   CDS             34792..51099
FT                   /*tag= b
FT                   /product= "NysJ protein"
FT   CDS             51155..57355
FT                   /*tag= c
FT                   /product= "NysK protein"
FT   CDS             57503..58687
FT                   /*tag= d
FT                   /product= "NysL protein"
FT   CDS             complement (58786..58980)
FT                   /*tag= e
FT                   /product= "NysM protein"
FT                   /note= "CDS does not include start codon"
FT   CDS             complement (59045..60241)
FT                   /*tag= f
FT                   /product= "NysN protein"
FT                   /note= "CDS does not include start codon"
FT   CDS             complement (60238..61296)
FT                   /*tag= g
FT                   /product= "NysD2 complete protein"
FT   CDS             120628..121308
FT                   /*tag= h
FT                   /product= "NysR4 (long) protein"
XX
PN   WO200159126-A2.
XX
PD   16-AUG-2001.
XX
PF   08-FEB-2001; 2001WO-GB00509.
XX
PR   08-FEB-2000; 2000GB-0002840.
PR   10-APR-2000; 2000GB-0008786.
PR   14-APR-2000; 2000GB-0009387.
XX
PA   (UYNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.
PA   (SNTF ) SINTEF STIFTELSEN IND TEK FORSK.
PA   (ALPH-) ALPHARMA AS.
PA   (SINV-) SINVENT AS.
PA   (DZIE/) DZIEGLEWSKA H.
PA   (ZOTC/) ZOTCHEV S B.
PA   (SEKU/) SEKUROVA O N.
PA   (FJAE/) FJAERVIK E.
PA   (BRAU/) BRAUTASET T.
PA   (STRO/) STROM A R.
XX
PI   Zotchev SB,  Sekurova ON,  Fjaervik E,  Brautaset T,  Strom AR;
PI   Valla S,  Ellingsen TE,  Sletta H,  Gulliksen O;
XX
DR   WPI; 2001-557614/62.
DR   P-PSDB; AAE10143, AAE10144, AAE10145, AAE10146, AAE10147, AAE10148,
DR   AAE10149, AAE10150.
XX
PT   New nystatin polyketide synthase polynucleotides and polypeptides,
PT   useful as antibiotics and antifungals -
XX
PS   Claim 1; Page 188-254; 266pp; English.
XX
CC   The present invention relates to the cloning and sequencing of the gene
CC   cluster encoding a modular type I polyketide synthase (PKS) enzyme
CC   involved in the biosynthesis of the macrolide antibiotic nystatin.
CC   The nystatin PKS is useful as antifungal antibiotics. The present
CC   sequence is a Streptomyces noursei nystatin PKS gene cluster DNA.
XX
SQ   Sequence 125401 BP; 15664 A; 49692 C; 42871 G; 17174 T; 0 other;


  Query Match             2.8%;  Score 38.2;  DB 22;  Length 125401;
  Best Local Similarity   45.6%;  Pred. No. 3.5;
  Matches  118;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy    193 ggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccatcgcggcgaaaag 252
          |||| || | |   ||||   ||   | ||     || | ||  | |   || |    |
Db  66232 ggcgcggagatcttcgcggacgtggaatgcccgcccggcaccgccgacgacgcccccgac 66291

Qy    253 cgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatcgagttcgaacgc 312
          | || ||| ||  | | | || | | ||  |   |  ||| ||  || | | | ||  ||
Db  66292 cacggactgcaccccgccctgctcgacgcggcccggcacgccgccatggcggtggacggc 66351

Qy    313 tacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngcaatctgatccaa 372
            || |||||  || | |     | |  |     | |   |  |   ||    | | |
Db  66352 accgtgcccgtcgcctggcacggcgtccggctgcacgccgtcggcgccaccgcgctgcgg 66411

Qy    373 tccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccctgccgtcgatgcc 432
            | |  |  |  |  |  |     |  | |  |  ||        | |||  | | |
Db  66412 gtccgcatccgccccaccacgaccggcacgctgaccctcaccgcggtcgacgtgcacggc 66471

Qy    433 gagccgttcgccatcttcg 451
          | |||| ||| || | |||
Db  66472 gcgccggtcgtcaccgtcg 66490


RESULT  19
AAV34766
ID   AAV34766 standard; RNA; 9759 BP.
XX
AC   AAV34766;
XX
DT   11-SEP-1998  (first entry)
XX
DE   Rubella virus RA27/3 genomic sequence.
XX
KW   genomic sequence; viral genome; rubella; RA27/3; vaccine; CRS;
KW   congenital rubella syndrome; vaccination; ss.
XX
OS   Rubella virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             41..6388
FT                   /*tag= a
FT                   /product= NSP-ORF, nonstructural proteins
FT   CDS             6509..9700
FT                   /*tag= b
FT                   /product= SP-ORF, virion structural proteins C,E2,E1
XX
PN   WO9820901-A1.
XX
PD   22-MAY-1998.
XX
PF   07-NOV-1997;   97WO-US20399.
XX
PR   12-NOV-1996;   96US-0030734.
XX
PA   (UYGE-) UNIV GEORGIA STATE RES FOUND.
XX
PI   Frey TK,  Pougatchev KV;
XX
DR   WPI; 1998-297616/26.
DR   P-PSDB; AAW59276, AAW59277.
XX
PT   Genomic sequence for RA27/3 strain of rubella virus - useful to
PT   identify further attenuated strains of rubella virus, e.g. for use
PT   in vaccines for congenital rubella syndrome
XX
PS   Claim 3; Fig 1; 34pp; English.
XX
CC   The sequence is that of the rubella RA27/3 genome.
CC   The nucleic acid can be used to identify further attenuated
CC   strains of rubella virus, by sequencing putative strains and determining
CC   whether specific attenuating mutation sites are present by comparison
CC   with the sequence of RA27/3. Such strains, and the nucleic acid
CC   sequence, are useful in vaccines against rubella virus, an
CC   important human pathogen. Whilst acute infection is usually benign,
CC   infection during early pregnancy can result in viral passage across the
CC   placenta and replication in the foetus, causing severe birth defects
```

```
CC   (congenital rubella syndrome, CRS). Eight live attenuated rubella virus
CC   vaccines to prevent CRS have previously been developed by serial passage
CC   through culture, and vaccine comprising the RA27/3 strain is the most
CC   commonly used.  Genetic characterisation of RA27/3 is also important to
CC   assure that the sequence (and therefore the attenuated phenotype) is
CC   maintained during vaccine production, and to determine whether
CC   complications sometimes occurring after vaccination (e.g. chronic
CC   arthritis) are of wild-type or vaccine origin.  The nucleic acid may
CC   also be used to identify specific attenuating mutation sites, by
CC   introducing RA27/3-specific nucleotides into a wild-type infectious
CC   clone (e.g. an existing infectious clone based on the Therien strain)
CC   and determining whether the resulting virus is attenuated. It may
CC   also be used to express RA27/3 in recombinant host cells, by
CC   transfecting host cells with the expression vector containing
CC   the nucleic acid and culturing the cells under suitable conditions.
XX
SQ   Sequence 9759 BP; 1458 A; 3775 C; 3008 G; 1518 T; 0 other;


  Query Match             2.8%;  Score 37.2;  DB 19;  Length 9759;
  Best Local Similarity   45.7%;  Pred. No. 2;
  Matches  123;  Conservative    0;  Mismatches  146;  Indels    0;  Gaps    0;

Qy      55 gtcatttatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggt 114
           | ||  | || || | ||| |         | |  |  |     ||| .||    |  |
Db    3944 ggcacctgtgccgcgaccgacgaggggctggcccaggcgtactacgacgacctcgaggtg 4003

Qy     115 atgcgccccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttt 174
             ||||  |  |    | | || ||   | |  ||| ||  |  || ||||   |
Db    4004 cgccgcctcggggatgacgccatggcccgggcggccctcgcatcagtccaacgccctcgc 4063

Qy     175 gaagacaaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgcc 234
            ||| |    |  ||    | ||   |   |  ||  | |   | |||||  | | ||  |
Db    4064 aaaggcccttacaatatcagggtatggaacatggccgcaggcgctggcaagaccacccgc 4123

Qy     235 atccatcgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgac 294
           ||||     | |   | ||||| |   |  |  |||       ||  ||    |   ||
Db    4124 atcctcgctgccttcacgcgcgaagacctttacgtctgccccaccaatgcgctcctgcac 4183

Qy     295 gaaatcgagttcgaacgctacgcgcccga 323
           || ||| ||  | ||| |  ||||| |||
Db    4184 gagatccaggccaaactccgcgcgcgcga 4212


RESULT  20
AAZ09101/c
ID   AAZ09101 standard; DNA; 763 BP.
XX
AC   AAZ09101;
XX
DT   18-OCT-1999  (first entry)
XX
DE   Mistletoe lectin A-chain MLA DNA consensus sequence.
XX
KW   Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
KW   ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW   lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW   cancer; cytotoxicity; antigen; isoform; ss.
XX
OS   Viscum album.
XX
PN   DE19804210-A1.
XX
PD   12-AUG-1999.
XX
PF   03-FEB-1998;   98DE-1004210.
XX
PR   03-FEB-1998;   98DE-1004210.
XX
PA   (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
PI   Morris P,  Stiefel T,  Voelter W,  Welters P;
XX
DR   WPI; 1999-445335/38.
XX
PT   Preparation of mistletoe lectins in heterologous systems,
PT   particularly for use as anticancer agents and immunostimulants
XX
PS   Claim 12; Page 34; 78pp; German.
XX
CC   This invention describes a novel mistletoe lectin (I) and its fragments
CC   which have antitumour and immunostimulatory activity. The A-chain (MLA)
CC   of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC   ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC   lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC   fragments are used to treat uncontrolled cell growth (particularly
CC   cancers) and if they lack cytotoxicity, to increase the strength of the
CC   immune response, particularly to a co-administered antigen
CC   (tumour-associated, bacterial or viral). The method allows production of
CC   mistletoe lectin, and its individual chains, in many different isoforms
CC   and on a large scale, at any time of the year. Recombinant products are
CC   free from toxins present in natural mistletoe extracts. This sequence
CC   represents a consensus sequence of mistletoe lectin A-chain (MLA) DNA
CC   described in the specification.
XX
SQ   Sequence 763 BP; 163 A; 215 C; 175 G; 170 T; 40 other;


  Query Match             2.8%;  Score 37;  DB 20;  Length 763;
  Best Local Similarity   47.6%;  Pred. No. 0.65;
  Matches   79;  Conservative    8;  Mismatches   79;  Indels    0;  Gaps    0;

Qy     871 gaattggttgacgcagacaaccgcgtgatttccggttcggtattgaacggcgcgattaca 930
           || | || ||  || |   |    |||  || | |  | |  | |  || ||||   | |
Db     321 GANTCGGGTGGTGCCGGTGAAGAGGTGCGTTTCCGCGCCGCGTGGTGCGTCGCGCAAAAA 262

Qy     931 caaggcgcgcacgattatttgggacgctaccacaatcagatttccgttatcgaagaaggc 990
             |||   |  ||  |  |||| | || || ||   :::|||       |||| |  |||
Db     261 GTAGGATTGGTCGCCTGCTTGGTAAGCCACGACGTAMRSATTGGTAACGTCGATGGCGGC 202

Qy     991 cgcagcaaagagctgttcggctgggttgcgccgcagccggacaaat 1036
           || : :  ||  :::   |      | || ||  || |  |||||||
Db     201 CGTRAYCGAGTCYYSCCCCTGGTTGGTGAGCTCCACCAAGACAAAT 156


RESULT  21
AAZ09100/c
ID   AAZ09100 standard; DNA; 1598 BP.
XX
AC   AAZ09100;
XX
DT   18-OCT-1999  (first entry)
XX
DE   Mistletoe lectin DNA consensus sequence.
XX
KW   Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
KW   ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW   lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW   cancer; cytotoxicity; antigen; isoform; ss.
XX
OS   Viscum album.
XX
PN   DE19804210-A1.
XX
PD   12-AUG-1999.
XX
PF   03-FEB-1998;   98DE-1004210.
XX
PR   03-FEB-1998;   98DE-1004210.
XX
PA   (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
PI   Morris P,  Stiefel T,  Voelter W,  Welters P;
XX
DR   WPI; 1999-445335/38.
XX
```

```
PT   Preparation of mistletoe lectins in heterologous systems,
PT   particularly for use as anticancer agents and immunostimulants
XX
PS   Claim 11; Page 32-33; 78pp; German.
XX
CC   This invention describes a novel mistletoe lectin (I) and its fragments
CC   which have antitumour and immunostimulatory activity. The A-chain (MLA)
CC   of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC   ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC   lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC   fragments are used to treat uncontrolled cell growth (particularly
CC   cancers) and if they lack cytotoxicity, to increase the strength of the
CC   immune response, particularly to a co-administered antigen
CC   (tumour-associated, bacterial or viral). The method allows production of
CC   mistletoe lectin, and its individual chains, in many different isoforms
CC   and on a large scale, at any time of the year. Recombinant products are
CC   free from toxins present in natural mistletoe extracts. This sequence
CC   represents a consensus sequence of mistletoe lectin DNA described in the
CC   specification.
XX
SQ   Sequence 1598 BP; 373 A; 396 C; 392 G; 356 T; 81 other;

  Query Match             2.8%;  Score 37;  DB 20;  Length 1598;
  Best Local Similarity   47.6%;  Pred. No. 0.94;
  Matches   79;  Conservative    8;  Mismatches   79;  Indels    0;  Gaps    0;

Qy     871 gaattggttgacgcagacaaccgcgtgatttccggttcggtattgaacggcgcgattaca 930
           || | || ||  || |   |    |||  || | |  | |  | |  || ||||   | |
Db     321 GANTCGGGTGGTGCCGGTGAAGAGGTGCGTTTCCGCGCCGCGTGGTGCGTCGCGCAAAAA 262

Qy     931 caaggcgcgcacgattatttgggacgctaccacaatcagatttccgttatcgaagaaggc 990
             |||   |  ||  |  |||| | || || ||   :::|||       |||| |  |||
Db     261 GTAGGATTGGTCGCCTGCTTGGTAAGCCACGACGTAMRSATTGGTAACGTCGATGGCGGC 202

Qy     991 cgcagcaaagagctgttcggctgggttgcgccgcagccggacaaat 1036
           || : :  ||  :::  |      | || ||  || |  |||||||
Db     201 CGTRAYCGAGTCYYSCCCCTGGTTGGTGAGCTCCACCAAGACAAAT 156


RESULT  22
AAZ95378
ID   AAZ95378 standard; DNA; 1401 BP.
XX
AC   AAZ95378;
XX
DT   01-JUN-2000  (first entry)
XX
DE   Chlamydia pneumoniae antigen CPN100605 protein encoding DNA SEQ ID NO:1.
XX
KW   Chlamydia pneumoniae; antigen; CPN100605 protein; immunisation;
KW   vaccine; infection; antibacterial; antiinflammatory; bronchitis;
KW   community acquired pneumonia; upper respiratory tract infection;
KW   sinusitis; ds.
XX
OS   Chlamydia pneumoniae.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1401
FT                   /*tag=  a
FT                   /product= "Antigen CPN100605 protein"
FT                   /note= "no stop codon given"
XX
PN   WO200006742-A2.
XX
PD   10-FEB-2000.
XX
PF   27-JUL-1999;   99WO-IB01331.
XX
PR   27-JUL-1998;   98US-0094195.
PR   26-JUL-1999;   99US-0361443.
XX
PA   (CONN-) CONNAUGHT LAB LTD.
XX
PI   Murdin AD,  Oomen RP;
XX
DR   WPI; 2000-205466/18.
DR   P-PSDB; AAY82082.
XX
PT   Chlamydia pneumoniae antigens used for immunization and protection
PT   against Chlamydia diseases -
XX
PS   Claim 1; Fig 1; 48pp; English.
XX
CC   The present sequence encodes the Chlamydia pneumoniae antigen CPN100605
CC   protein. The CPN100605 protein has antibacterial and antiinflammatory
CC   activities. The Chlamydia pneumoniae CPN100605 polynucleotide and protein
CC   can be used in vaccination methods for preventing and treating Chlamydia
CC   infection (e.g. infections caused by C. trachomatis, C. psittaci,
CC   C. pneumoniae or C. pecorum). The polynucleotide can be used to produce
CC   the protein recombinantly, in the construction of vaccine vectors,
CC   as a vaccine agent, and in the construction of an attenuated Chlamydia
CC   strain. The protein are also be useful as a vaccine agent, and for the
CC   preparation of medicaments for treating or preventing Chlamydia
CC   infection, e.g. community acquired pneumonia, and upper respiratory
CC   tract infections such as bronchitis and sinusitis.
XX
SQ   Sequence 1401 BP; 438 A; 308 C; 239 G; 416 T; 0 other;

  Query Match             2.7%;  Score 36.6;  DB 21;  Length 1401;
  Best Local Similarity   52.3%;  Pred. No. 1.2;
  Matches   78;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy     128 tgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagacaaaaagn 187
           | || ||  || |||| ||||| |||      |||         |    ||| ||||  |
Db     134 ttaaggtagagcaaggagatgcggtctgttcaggagctcctatagcagaatacaagcact 193

Qy     188 atccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccatcgcggcg 247
            |||   |       ||||| | |  ||||||||   | |  |||| || |  || ||
Db     194 ttcctaacacctacattacctctcacgtttcaggagtagttaccgctatacgacgtggaa 253

Qy     248 aaaagcgcgtacttcagtcggtcgtgatt 276
           | || ||    |||      ||| | |||
Db     254 ataaacgttctcttttagatgtcatcatt 282


RESULT  23
AAD08837
ID   AAD08837 standard; cDNA; 1325 BP.
XX
AC   AAD08837;
XX
DT   04-SEP-2001  (first entry)
XX
DE   Human G-protein coupled receptor-3 (GCREC-3) cDNA.
XX
KW   Human; G-protein coupled receptor-3; GCREC-3; gene therapy; cirrhosis;
KW   transgenic animal; proliferative disorder; actinic keratosis; hepatitis
KW   nephrotropic; cancer; breast; bladder; bone marrow; brain; uterus;
KW   leukaemia; adenocarcinoma; lymphoma; melanoma; myeloma; epilepsy; stroke;
KW   neurological disorder; Alzheimer's disease; Parkinson's disease; nausea;
KW   Huntington's disease; multiple sclerosis; dementia; angina pectoris;
KW   central nervous system disorder; cardiovascular disorder; hypertension;
KW   atherosclerosis; congestive heart failure; gastrointestinal disorder;
KW   dysphagia; peptic oesophagitis; spasm; gastritis; anorexia; pyrosis;
KW   pancreatitis; Crohn's disease; diarrhoea; autoimmune disorder; anaemia;
KW   inflammatory disorder; Acquired Immune Deficiency Syndrome; AIDS;
KW   Addison's diseases; allergy; asthma; diabetes mellitus; antithyroid;
KW   atopic dermatitis; glomerulonephritis; Grave's disease; psoriasis;
KW   rheumatoid arthritis; ulcerative colitis; osteoporosis; antifungal;
KW   metabolic disorder; obesity; nootropic; protozoacide; virucide; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
```

```
FT   CDS             42..1310
FT                   /*tag= a
FT                   /product= "Human GCREC-3 protein"
XX
PN   WO200142288-A2.
XX
PD   14-JUN-2001.
XX
PF   07-DEC-2000; 2000WO-US33382.
XX
PR   10-DEC-1999;   99US-0172852.
PR   22-DEC-1999;   99US-0171732.
PR   14-JAN-2000; 2000US-0176148.
PR   21-JAN-2000; 2000US-0177331.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Burford N,  Baughn MR,  Au-Young J,  Yang J,  Lu DAM,  Reddy R;
XX
DR   WPI; 2001-381635/40.
DR   P-PSDB; AAE04547.
XX
PT   New human G-protein coupled receptor polypeptides for diagnosing,
PT   preventing, and treating cell proliferative, neurological,
PT   cardiovascular, gastrointestinal, autoimmune and metabolic disorders -
XX
PS   Claim 5; Page 160-161; 175pp; English.
XX
CC   The present sequence is human G-protein coupled receptor-3 (GCREC-3)
CC   cDNA. GCREC is useful in somatic or germline gene therapy to correct a
CC   genetic deficiency, to express a conditionally lethal gene product and
CC   to express a protein which affords protection against intracellular
CC   parasites and also for diagnosis of disorders associated with expression
CC   of GCREC. GCREC is also useful for generating hybridisation probes useful
CC   in mapping the naturally occurring genomic sequences and to create
CC   knockin humanised animals (pigs) or transgenic animals (mice or rats) to
CC   model human diseases. GCREC is used to diagnose, prevent and treat
CC   proliferative disorders (actinic keratosis, arteriosclerosis, cirrhosis,
CC   hepatitis and cancer); cancer (breast, bladder, bone marrow, brain,
CC   uterus cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma)
CC   neurological disorders (epilepsy, stroke, Alzheimer's, Huntington's,
CC   Parkinson's disease, multiple sclerosis, dementia and other central
CC   nervous system disorders); cardiovascular disorders (angina pectoris,
CC   hypertension, atherosclerosis, congestive heart failure);
CC   gastrointestinal disorders (dysphagia, peptic oesophagitis, oesophageal
CC   spasm, gastritis, gastric carcinoma, anorexia, nausea, abdominal angina,
CC   pyrosis, pancreatitis, Crohn's disease, diarrhoea); autoimmune/
CC   inflammatory disorders (acquired immunodeficiency syndrome (AIDS),
CC   Addison's diseases, allergies, anaemia, asthma, diabetes mellitus, atopic
CC   dermatitis, glomerulonephritis, Grave's disease, osteoarthritis,
CC   psoriasis, rheumatoid arthritis, ulcerative colitis, bacterial, fungal,
CC   parasitic, protozoal and helminthic infections) and metabolic disorders
CC   (obesity, osteoporosis, viral infections).
XX
SQ   Sequence 1325 BP; 195 A; 473 C; 403 G; 254 T; 0 other;

  Query Match             2.7%;  Score 36.4;  DB 22;  Length 1325;
  Best Local Similarity   51.9%;  Pred. No. 1.3;
  Matches   82;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy    973 tccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcgccgcagccggac 1032
          | | ||| ||  || |||  | ||||    |||  ||||  ||   | |||||| | || |
Db    933 tactttagcgccgagggcttccgcaacaccctgcgcggcctgggcactccgcaccgggcc 992

Qy   1033 aaatactccatcacgcgtacgaccctcggccatttcctgaaaaacaaactcttcaagttc 1092
          |    |||   |||      ||| |  |      ||  | ||  | ||   | |  ||
Db    993 aggacctcggccaccaacgggacgcgggcggcgctcgcgcaatccgaaaggtccgccgtc 1052

Qy   1093 acgacagccgtcaacggtggcgaccgcgccatggtgcc 1130
          || || | || || | |   ||  |||||||  ||
Db   1053 accaccgacgccaccaggccggatgccgccatgtcccc 1090
```

```
RESULT  24
AAF87442/c
ID   AAF87442 standard; DNA; 4096 BP.
XX
AC   AAF87442;
XX
DT   09-JUL-2001  (first entry)
XX
DE   Corynebacterium thermoaminogenes odhA nucleotide sequence.
XX
KW   Corynebacterium; thermophilic; amino acid biosynthesis; enzyme;
KW   thermotolerant; aceA; accBC; dtsR1; dtsR2; pfk; scrB; gluABCD;
KW   pdhA; pc; ppc; acn; icd; lpd; odhA; ds.
XX
OS   Corynebacterium thermoaminogenes.
XX
FH   Key             Location/Qualifiers
FT   CDS             250..3954
FT                   /*tag= a
FT                   /partial
FT                   /product= "odhA protein"
FT                   /note= "this sequence does not contain a start codon"
XX
PN   WO200125447-A1.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-JP06913.
XX
PR   04-OCT-1999;   99JP-0282716.
PR   01-NOV-1999;   99JP-0311147.
PR   21-APR-2000; 2000JP-0120687.
XX
PA   (AJIN ) AJINOMOTO CO INC.
XX
PI   Hirano S,  Nonaka G,  Matsuzaki Y,  Akiyoshi N,  Nakamura K,  Kimura E;
PI   Osumi T,  Matsui K,  Kawahara Y,  Kurahashi O,  Nakamatsu T;
PI   Sugimoto S;
XX
DR   WPI; 2001-300170/31.
DR   P-PSDB; AAB83185.
XX
PT   Proteins and their DNA useful for microbial production of L-amino acids
PT   -
XX
PS   Claim 44; Page 161-166; 215pp; Japanese.
XX
CC   The present sequence is provided in a specification relating to genes
CC   encoding thermophilic amino acid biosynthesis system enzymes of
CC   the thermotolerant bacterium Corynebacterium thermoaminogenes.
CC   The novel proteins retain at least 30% isocitrate ligase activity
CC   after heating at 50oC for 5 minutes. DNA fragments encoding the
CC   enzymes were isolated from a Corynebacterium thermoaminogenes
CC   chromosomal DNA plasmid library by PCR. The DNA may be used for
CC   developing strains of amino acid producing microorganisms.
XX
SQ   Sequence 4096 BP; 824 A; 1432 C; 1143 G; 697 T; 0 other;

  Query Match             2.7%;  Score 36.4;  DB 22;  Length 4096;
  Best Local Similarity   45.7%;  Pred. No. 2.3;
  Matches  121;  Conservative    0;  Mismatches  144;  Indels    0;  Gaps    0;

Qy     60 ttatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcg 119
          | | | || |    ||  |  ||| ||| || || | | || |   |||||||  ||
Db   1534 TCAGGCCGTGGGTCTCGATGTCGAGGTCACGGTGATCCGGGACGGGCATGCCGGGCTGGA 1475

Qy    120 cccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaaga 179
          |||   |  |   ||    | |||||||  || |     ||    ||     | || ||
Db   1474 CCCAGGGCAGTGGGTTGGTGTCGGCGATGAGGTGACCGCGGGAGCGGTAGGCCTCGATGA 1415

Qy    180 caaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatcca 239
```

```
Db   1414 GCTGCATGACACGGGTGTTCTTGTCCACACCGGTGTTGGGCAGGTCCTGTGCCCAGCGCA 1355

Qy    240 tcgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaat 299
          ||| || | |  |  ||| | ||     |  |   |  |   ||| ||| || | |
Db   1354 TCGGGGTGTAGGGAACGTTCATCTCCTCGAAGATGTGATCCCAGAACGCATCGTCCACGA 1295

Qy    300 cgagttcgaacgctacgcgcccgaa 324
            || | | ||     ||||  |||
Db   1294 GCAGCTGGGACATGGTGCGCAGGAA 1270


RESULT 25
AAZ09104/c
ID   AAZ09104 standard; DNA; 762 BP.
XX
AC   AAZ09104;
XX
DT   18-OCT-1999 (first entry)
XX
DE   Mistletoe lectin A1 DNA fragment.
XX
KW   Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
KW   ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW   lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW   cancer; cytotoxicity; antigen; isoform; lectin A1; ds.
XX
OS   Viscum album.
XX
PN   DE19804210-A1.
XX
PD   12-AUG-1999.
XX
PF   03-FEB-1998;   98DE-1004210.
XX
PR   03-FEB-1998;   98DE-1004210.
XX
PA   (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
PI   Morris P,  Stiefel T,  Voelter W,  Welters P;
XX
DR   WPI; 1999-445335/38.
DR   P-PSDB; AAY25980.
XX
PT   Preparation of mistletoe lectins in heterologous systems,
PT   particularly for use as anticancer agents and immunostimulants
XX
PS   Claim 15; Fig 2A; 78pp; German.
XX
CC   This invention describes a novel mistletoe lectin (I) and its fragments
CC   which have antitumour and immunostimulatory activity. The A-chain (MLA)
CC   of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC   ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC   lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC   fragments are used to treat uncontrolled cell growth (particularly
CC   cancers) and if they lack cytotoxicity, to increase the strength of the
CC   immune response, particularly to a co-administered antigen
CC   (tumour-associated, bacterial or viral). The method allows production of
CC   mistletoe lectin, and its individual chains, in many different isoforms
CC   and on a large scale, at any time of the year. Recombinant products are
CC   free from toxins present in natural mistletoe extracts. This sequence
CC   encodes a fragment of the mistletoe lectin A1 protein.
XX
SQ   Sequence 762 BP; 171 A; 230 C; 183 G; 178 T; 0 other;


  Query Match             2.7%;  Score 36;  DB 20;  Length 762;
  Best Local Similarity   50.6%;  Pred. No. 1.3;
  Matches   87;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy    865 gcgggcgaattggttgacgcagacaaccgcgtgatttccggttcggtattgaacggcgcg 924
          | | | | || || ||  || |   |    ||| || | |  | |  | |  || ||||
Db    327 GAGAGAGGATCGGGTGGTGCCGGTGAAGAGGTGCGTTTCCGCGCCGCGTGGTGCGTCGCG 268

Qy    925 attacacaaggcgcgcacgattatttgggacgctaccacaatcagatttccgttatcgaa 984
          | |  |||   |  ||  |  |||| | || || ||      |||      ||||
Db    267 CAAAAAGTAGGATTGGTCGCCTGCTTGGTAAGCCACGACGTAAGCATTGGTAACGTCGAT 208

Qy    985 gaaggccgcagcaaagagctgttcggctgggttgcgccgcagccggacaaat 1036
          |  ||||| | |  ||  |||  |      | || ||  || |  |||||||
Db    207 GGCGGCCGTAACCGAGTCCTGCCCCTGGTTGGTGAGCTCCACCAAGACAAAT 156


RESULT 26
AAH16235
ID   AAH16235 standard; cDNA; 3230 BP.
XX
AC   AAH16235;
XX
DT   26-JUN-2001 (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:15062.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 15062; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
```

```
SQ   Sequence 3230 BP; 674 A; 904 C; 929 G; 723 T; 0 other;

  Query Match             2.7%;  Score 36;  DB 22;  Length 3230;
  Best Local Similarity   58.3%;  Pred. No. 2.7;
  Matches   63;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy     608 atgctgccaacatcgaaacacatgaattcggcggcccgcatccggccggtttgagtggca 667
           ||||| || ||  |    ||  ||  | |||   ||||    || | |||||||| | |
Db    1425 atgctacccactcccctccatctgtgtgcgggaccccgtgagcgactggtttgagagcct 1484

Qy     668 cgcacattcatttcattgagccggtcggtgcaaacaaaaccgtttgga 715
           ||| || |   | |||||   ||  |||   || |||  ||  || ||
Db    1485 cgcccagtgcctgcattggaacgtccggaagaagcaagcccacttcga 1532


RESULT  27
AAA81473/c
ID   AAA81473 standard; DNA; 92934 BP.
XX
AC   AAA81473;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_21 SEQ ID NO:21.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 471-498; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 92934 BP; 24051 A; 24249 C; 21902 G; 22730 T; 2 other;

  Query Match             2.7%;  Score 36;  DB 21;  Length 92934;
  Best Local Similarity   51.9%;  Pred. No. 14;
  Matches   81;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy     870 cgaattggttgacgcagacaaccgcgtgatttccggttcggtattgaacggcgcgattac 929
           ||||||||||  |     | | |||    |  ||      | | | |||| ||||||
Db   64768 CGAATTGGTTTCCATCCTCGAACGCCCTGTAACCAAAGAAGAAATCAACGCCGCGATGAA 64709

Qy     930 acaaggcgcgcacgattatttgggacgctaccacaatcagatttccgttatcgaagaagg 989
           |   | |||   ||| | ||  ||   | || |  |||| ||      |  ||| |  |
Db   64708 AGCCGCCGCCAGCGAATCTTACGGCTACAACGAAGATCAAATCGTTTCTTCCGACGTTGT 64649

Qy     990 ccgcagcaaagagctgttcggctgggttgcgccgca 1025
           | ||| | || |    |    ||  | |||| | ||
Db   64648 CGGCATCGAATACGGCTCGCTCTTCGATGCGACCCA 64613


RESULT  28
AAF21613/c
ID   AAF21613 standard; DNA; 172325 BP.
XX
AC   AAF21613;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Neisseria meningitidis B nucleotide sequence SEQ ID NO:114.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW   diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW   ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200066791-A1.
XX
PD   09-NOV-2000.
XX
PF   08-MAR-2000; 2000WO-US05928.
XX
PR   30-APR-1999;   99US-0132068.
PR   08-OCT-1999;   99WO-US23573.
PR   28-FEB-2000; 2000GB-0004695.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Pizza M,  Hickey E,  Peterson J,  Tettelin H,  Venter JC,  Masignani V;
PI   Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V,  Rappuoli R;
PI   Frazer CM,  Grandi G;
XX
DR   WPI; 2000-647603/62.
XX
PT   Neisseria meningitidis B full length genome sequence and open reading
PT   frames are used to detect, treat and prevent Neisserial infections -
XX
PS   Claim 7; Appendix A; 692pp; English.
XX
CC   The present invention describes the full length genome of
CC   Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC   to AAF21613 represent fragments of the NMB genomic sequence, as the
CC   sequence was too long to go in a record on its own if was split into 8
CC   sequences which overlap each other at the beginning and end of each
```

CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.
XX
SQ Sequence 172325 BP; 43072 A; 47583 C; 41465 G; 40205 T; 0 other;

Query Match 2.7%; Score 36; DB 21; Length 172325;
Best Local Similarity 51.9%; Pred. No. 20;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 870 cgaattggttgacgcagacaaccgcgtgatttccggttcggtattgaacggcgcgattac 929
Db 168435 CGAATTGGTTTCCATCCTCGAACGCCCTGTAACCAAAGAAGAAATCAACGCCGCGATGAA 168376

Qy 930 acaaggcgcgcacgattatttgggacgctaccacaatcagatttccgttatcgaagaagg 989
Db 168375 AGCCGCCGCCAGCGAATCTTACGGCTACAACGAAGATCAAATCGTTTCTTCCGACGTTGT 168316

Qy 990 ccgcagcaaagagctgttcggctgggttgcgccgca 1025
Db 168315 CGGCATCGAATACGGCTCGCTCTTCGATGCGACCCA 168280

RESULT 29
AAA81489/c
ID AAA81489 standard; DNA; 837096 BP.
XX
AC AAA81489;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_37 SEQ ID NO:37.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Claim 7; Page 629-865; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neissariae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
SQ Sequence 837096 BP; 207534 A; 227065 C; 205215 G; 197280 T; 2 other;

Query Match 2.7%; Score 36; DB 21; Length 837096;
Best Local Similarity 51.9%; Pred. No. 43;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 870 cgaattggttgacgcagacaaccgcgtgatttccggttcggtattgaacggcgcgattac 929
Db 833206 CGAATTGGTTTCCATCCTCGAACGCCCTGTAACCAAAGAAGAAATCAACGCCGCGATGAA 833147

Qy 930 acaaggcgcgcacgattatttgggacgctaccacaatcagatttccgttatcgaagaagg 989
Db 833146 AGCCGCCGCCAGCGAATCTTACGGCTACAACGAAGATCAAATCGTTTCTTCCGACGTTGT 833087

Qy 990 ccgcagcaaagagctgttcggctgggttgcgccgca 1025
Db 833086 CGGCATCGAATACGGCTCGCTCTTCGATGCGACCCA 833051

RESULT 30
AAC97005
ID AAC97005 standard; DNA; 699 BP.
XX
AC AAC97005;
XX
DT 23-FEB-2001 (first entry)
XX
DE Mycobacterium tuberculosis secreted protein cDNA #27.
XX
KW Mycobacterium tuberculosis secreted protein; MTSP; vaccine; ss.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200066143-A1.
XX
PD 09-NOV-2000.
XX
PF 04-MAY-2000; 2000WO-US12197.
XX
PR 04-MAY-1999; 99US-0132479.
PR 04-MAY-1999; 99US-0132503.
XX

```
PA   (PUBL-) PUBLIC HEALTH RES INST NEW YORK.
XX
PI   Gennaro ML,  Gomez MJ;
XX
DR   WPI; 2001-007151/01.
XX
PT   Novel Mycobacterium tuberculosis secreted polypeptides and
PT   polynucleotides useful in diagnosis, treatment and prophylaxis of
PT   tuberculosis  -
XX
PS   Claim 1; Fig 2; 60pp; English.
XX
CC   The present invention relates to Mycobacterium tuberculosis secreted
CC   proteins (MTSP), where the polypeptide has M. tuberculosis specific
CC   antigenic and immunogenic properties. Compositions of the invention may
CC   be useful for diagnosing Mycobacterium tuberculosis infection and as a
CC   vaccine against M. tuberculosis infection.
XX
SQ   Sequence 699 BP; 120 A; 232 C; 234 G; 113 T; 0 other;


  Query Match             2.7%;  Score 35.8;  DB 22;  Length 699;
  Best Local Similarity   55.1%;  Pred. No. 1.5;
  Matches   70;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy   1044 cacgcgtacgaccctcggccatttcctgaaaaacaaactcttcaagttcacgacagccgt 1103
            ||| |  || ||| ||||||||  |  |||     | ||| ||   |||||  | |||
Db    147 caccctaaccaccatcggccatgacaagaacggtgacctcatcggcttcacttccgccca 206

Qy   1104 caacggtggcgaccgcgccatggtgccgattggtacttacgagcgcgtaatgccgctaga 1163
            |   || |||    ||||   | |  |   ||    |  ||||  ||     ||| |||
Db    207 ctgtgggggcccgggcgcgcagatcgccgctgagggtgccgagaacgcgggcccggtagg 266

Qy   1164 catcctg 1170
            |||| ||
Db    267 catcatg 273


RESULT  31
AAI99683/c
ID   AAI99683 standard; DNA; 4403765 BP.
XX
AC   AAI99683;
XX
DT   15-JAN-2002  (first entry)
XX
DE   Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
KW   Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW   variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS   Mycobacterium tuberculosis.
XX
PN   US6294328-B1.
XX
PD   25-SEP-2001.
XX
PF   24-JUN-1998;   98US-0103840.
XX
PR   24-JUN-1998;   98US-0103840.
XX
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fleischmann RD,  White OR,  Fraser CM,  Venter JC;
XX
DR   WPI; 2001-647261/74.
XX
PT   Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT   determining the nucleotide sequence of the strain at positions in the
PT   genome corresponding to positions where M. tuberculosis strains CDC
PT   1551 and H37Rv differ  -
XX
PS   Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
CC   The invention relates to evaluating strain variation within and between
CC   different populations of the tuberculosis bacterial pathogen,
CC   Mycobacterium tuberculosis or related Mycobacterium by determining the
CC   nucleotide sequence of the first strain at positions in the complete
CC   sequence of the genome that correspond to positions that differ in the
CC   nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC   H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC   M. tuberculosis and has valuable application in the fields of
CC   tuberculosis genetics, epidemiology, patient treatment and epidemic
CC   monitoring.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ   Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;


  Query Match             2.7%;  Score 35.8;  DB 22;  Length 4403765;
  Best Local Similarity   55.1%;  Pred. No. 98;
  Matches   70;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy   1044 cacgcgtacgaccctcggccatttcctgaaaaacaaactcttcaagttcacgacagccgt 1103
            ||| |  || ||| ||||||||  |  |||     | ||| ||   |||||  | |||
Db 4102529 CACCCTAACCACCATCGGCCATGACAAGAACGGTGACCTCATCGGCTTCACTTCCGCCCA 4102470

Qy   1104 caacggtggcgaccgcgccatggtgccgattggtacttacgagcgcgtaatgccgctaga 1163
            |   || |||    ||||   | |  |   ||    |  ||||  ||     ||| |||
Db 4102469 CTGTGGGGGCCCGGGCGCGCAGATCGCCGCTGAGGGTGCCGAGAACGCGGGCCCGGTAGG 4102410

Qy   1164 catcctg 1170
            |||| ||
Db 4102409 CATCATG 4102403


RESULT  32
AAI99682/c
ID   AAI99682 standard; DNA; 4411529 BP.
XX
AC   AAI99682;
XX
DT   15-JAN-2002  (first entry)
XX
DE   Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
KW   Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW   variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS   Mycobacterium tuberculosis.
XX
PN   US6294328-B1.
XX
PD   25-SEP-2001.
XX
PF   24-JUN-1998;   98US-0103840.
XX
PR   24-JUN-1998;   98US-0103840.
XX
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fleischmann RD,  White OR,  Fraser CM,  Venter JC;
XX
DR   WPI; 2001-647261/74.
XX
PT   Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT   determining the nucleotide sequence of the strain at positions in the
PT   genome corresponding to positions where M. tuberculosis strains CDC
PT   1551 and H37Rv differ  -
XX
PS   Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
CC   The invention relates to evaluating strain variation within and between
CC   different populations of the tuberculosis bacterial pathogen,
```

```
CC   Mycobacterium tuberculosis or related Mycobacterium by determining the
CC   nucleotide sequence of the first strain at positions in the complete
CC   sequence of the genome that correspond to positions that differ in the
CC   nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC   H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC   M. tuberculosis and has valuable application in the fields of
CC   tuberculosis genetics, epidemiology, patient treatment and epidemic
CC   monitoring.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ   Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;


  Query Match             2.7%;  Score 35.8;  DB 22;  Length 4411529;
  Best Local Similarity   55.1%;  Pred. No. 98;
  Matches   70;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy    1044 cacgcgtacgaccctcggccatttcctgaaaaacaaactcttcaagttcacgacagccgt 1103
           ||| |  || ||| |||||||  |  |||     | ||| ||   |||||  | |||
Db 4110332 CACCCTAACCACCATCGGCCATGACAAGAACGGTGACCTCATCGGCTTCACTTCCGCCCA 4110273

Qy    1104 caacggtggcgaccgcgccatggtgccgattggtacttacgagcgcgtaatgccgctaga 1163
           |   || |||    ||||   | |  |   ||    |  ||||  ||     ||| |||
Db 4110272 CTGTGGGGGCCCGGGCGCGCAGATCGCCGCTGAGGGTGCCGAGAACGCGGGCCCGGTAGG 4110213

Qy    1164 catcctg 1170
           |||| ||
Db 4110212 CATCATG 4110206


RESULT  33
AAA08045
ID   AAA08045 standard; cDNA; 2558 BP.
XX
AC   AAA08045;
XX
DT   19-JUN-2000  (first entry)
XX
DE   f-Therien rubella virus cDNA clone sequence SEQ ID NO:3.
XX
KW   Rubella virus; infection; infectious; chimeric construct; anti-viral;
KW   vaccine; birth defect; autoimmune disease; ss.
XX
OS   Rubella virus.
XX
PN   US6054573-A.
XX
PD   25-APR-2000.
XX
PF   02-SEP-1997;   97US-0999733.
XX
PR   28-JUN-1991;   91US-0722334.
PR   19-JUL-1993;   93US-0093453.
PR   02-JUN-1995;   95US-0459041.
XX
PA   (UYGE-) UNIV GEORGIA STATE.
XX
PI   Abernathy ES,  Pougatchev K,  Frey TK;
XX
DR   WPI; 2000-328366/28.
XX
PT   Highly infectious rubella virus clones useful for developing a rubella
PT   vaccine that can be safely administered to pregnant and older women
PT   without risk of birth defects and autoimmune disease  -
XX
PS   Claim 2; Column 21-24; 17pp; English.
XX
CC   The present invention describes nucleic acid molecules (AAA08044 and
CC   AAA08045) which are fragments of the f-Therien rubella virus genome.
CC   AAA08044 and AAA08045 are used to replace the corresponding fragments of
CC   infectious rubella virus cDNA clone with low specific infectivity
CC   referred to as Robo102 to create a chimeric construct with high
CC   specific infectivity, Robo302. The highly infectious rubella virus
CC   clones are useful as molecular biology tools for studying rubella virus
CC   and can be used for developing recombinant vaccines against rubella.
CC   The rubella vaccines developed can be safely administered to pregnant
CC   and older women without risk of birth defects or autoimmune disease.
XX
SQ   Sequence 2558 BP; 379 A; 991 C; 805 G; 383 T; 0 other;


  Query Match             2.6%;  Score 35.6;  DB 21;  Length 2558;
  Best Local Similarity   45.4%;  Pred. No. 3.2;
  Matches  122;  Conservative    0;  Mismatches  147;  Indels    0;  Gaps    0;

Qy     55 gtcatttatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggt 114
          | ||  | || || | ||| |          | |  |  |     ||| ||    |  |
Db   1142 ggcacctgtgccgcgaccgacgaggggctggcccaggcgtactacgacgacctcgaggtg 1201

Qy    115 atgcgcccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgttt 174
             |||| |  |     | | || ||   | |  ||| ||  |  || ||||   |
Db   1202 cgccgcctcggggatgacgccatggcccgggcggccctcgcatcagtccaacgccctcgc 1261

Qy    175 gaagacaaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgcc 234
           ||| |    |  ||    | ||   |   |  ||  | |   | ||||| |   ||  |
Db   1262 aaaggcccttacaatatcagggtatggaacatggccgcaggcgctggcaagactacccgc 1321

Qy    235 atccatcgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgac 294
           ||||     | |   | ||||| |   |  | |||       ||  ||   |    ||
Db   1322 atcctcgctgccttcacgcgcgaagacctttacgtctgccccaccaatgcgctcctgcac 1381

Qy    295 gaaatcgagttcgaacgctacgcgcccga 323
           || ||| ||  | ||| |  ||||| |||
Db   1382 gagatccaggccaaactccgcgcgcgcga 1410


RESULT  34
AAQ97686
ID   AAQ97686 standard; RNA; 9757 BP.
XX
AC   AAQ97686;
XX
DT   27-FEB-1996  (first entry)
XX
DE   Infectious rubella virus RNA.
XX
KW   Rubella; vaccine; mutant; epitope; virus; autoimmune disease;
KW   pregnancy; foetal infection; vector; plasmid; ss.
XX
OS   Rubella virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             41..6658
FT                   /*tag=  a
FT                   /product= N-terminal transcript.
FT   misc_difference 2261..2263
FT                   /*tag=  b
FT                   /transl_except= CGU encodes Ala.
FT   misc_difference 6605..6607
FT                   /*tag=  c
FT                   /transl_except= GUC encodes asparagine or aspartic acid.
FT   misc_difference 8460..8462
FT                   /*tag=  d
FT                   /transl_except= CUG encodes Proline.
FT   misc_difference 8463..8465
FT                   /*tag=  e
FT                   /transl_except= CUC encodes Cysteine.
FT   misc_difference 9075..9077
FT                   /*tag=  f
FT                   /transl_except= UGG encodes Methionine.
XX
PN   US5439814-A.
XX
PD   08-AUG-1995.
```

```
XX
PF   28-JUN-1991;   91US-0722334.
XX
PR   28-JUN-1991;   91US-0722334.
PR   19-JUL-1993;   93US-0093453.
XX
PA   (GEOR-) GEORGIA STATE RES FOUND INC.
XX
PI   Dominguez G,  Frey TK,  Wang C;
XX
DR   WPI; 1995-283097/37.
DR   P-PSDB; AAR79048, AAR79049.
XX
PT   New DNA encoding infectious rubella virus - esp. non-pathogenic
PT   mutant virus for use in vaccines having reduced side effects
XX
PS   Claim 2; Columns 9-18; 27pp; English.
XX
CC   Non-pathogenic mutants of the DNA corresponding to this sequence are
CC   useful in vaccines (which may include epitopes from other viruses).
CC   The mutant vaccines are less likely to cause foetal infections,
CC   autoimmune disease or neurological symptoms, so can be administered
CC   safely to older or pregnant women.  The mutant sequence is
CC   preferably present in a vector, especially a bacterial plasmid that
CC   allows replication of the sequence.
XX
SQ   Sequence 9757 BP; 1458 A; 3782 C; 3008 G; 1509 U; 0 other;

  Query Match             2.6%;  Score 35.6;  DB 16;  Length 9757;
  Best Local Similarity   41.6%;  Pred. No. 6.2;
  Matches  112;  Conservative   10;  Mismatches  147;  Indels    0;  Gaps    0;

Qy      55 gtcatttatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggt 114
           | ||  : :| || | ||| |         | |  |  |     ||| ||    |  |
Db    3941 ggcaccugugccgcgaccgacgaggggcuggcccaggcguacuacgacgaccucgaggug 4000

Qy     115 atgcgcccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttt 174
             |||| |  |    | | || ||   | |  ||| :|  |  || ||||   |
Db    4001 cgccgccucggggaugacgccauggcccgggcggcccucgcaucaguccaacgcccucgc 4060

Qy     175 gaagacaaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgcc 234
            ||| |    |  |:    |  |:   |   |  ||  | |    | ||||| | ·  ||  |
Db    4061 aaaggcccuuacaauaucaggguauggaacauggccgcaggcgcuggcaagacuacccgc 4120

Qy     235 atccatcgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgac 294
           |:||     | |   | |||||  |   |  :  |:|       ||  :|   ·|   ||
Db    4121 auccucgcugccuucacgcgcgaagaccuuuacgucugccccaccaaugcgcuccugcac 4180

Qy     295 gaaatcgagttcgaacgctacgcgcccga 323
           || |:| ||  | ||| |  ||||| |||
Db    4181 gagauccaggccaaacuccgcgcgcgcga 4209


RESULT  35
AAT89642
ID   AAT89642 standard; cDNA; 9759 BP.
XX
AC   AAT89642;
XX
DT   26-JAN-1998  (first entry)
XX
DE   Infectious rubella virus cDNA clone.
XX
KW   Rubella; German measles; foetal death; birth defect; vaccine;
KW   immunisation; non-pathogenic mutant; genetic engineering; ss.
XX
OS   Rubella virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             41..6388
FT                   /*tag=  a
FT                   /note= "encodes p150 and p90"
FT   CDS             6509..9700
FT                   /*tag=  b
FT                   /note= "encodes capsid, envelope 1 and envelope 2
FT                          glycoproteins"
XX
PN   US5663065-A.
XX
PD   02-SEP-1997.
XX
PF   28-JUN-1991;   91US-0722334.
XX
PR   02-JUN-1995;   95US-0459041.
PR   28-JUN-1991;   91US-0722334.
PR   19-JUL-1993;   93US-0093453.
XX
PA   (UYGE-) UNIV GEORGIA STATE RES FOUND.
XX
PI   Dominguez G,  Frey TK,  Wang C;
XX
DR   WPI; 1997-447930/41.
XX
PT   DNA encoding infectious rubella virus - useful for producing
PT   non-pathogenic mutants for use as vaccines
XX
PS   Claim 1; Column 9-18; 19pp; English.
XX
CC   AAT89642 is a cDNA molecule encoding an infectious rubella virus. The
CC   sequence can be modified to give non-pathogenic mutants of this clone.
CC   Such mutants are useful as vaccines for protecting against rubella
CC   infection. The non-pathogenic clones can be used for the vaccination
CC   of pregnant and older women with decreased risk of causing foetal
CC   infection, autoimmune disease or neurological symptoms.
XX
SQ   Sequence 9759 BP; 1458 A; 3784 C; 3007 G; 1510 T; 0 other;

  Query Match             2.6%;  Score 35.6;  DB 18;  Length 9759;
  Best Local Similarity   45.4%;  Pred. No. 6.2;
  Matches  122;  Conservative    0;  Mismatches  147;  Indels    0;  Gaps    0;

Qy      55 gtcatttatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggt 114
           | ||  | || || | ||| |         | |  |  |     ||| ||    |  |
Db    3944 ggcacctgtgccgcgaccgacgaggggctggcccaggcgtactacgacgacctcgaggtg 4003

Qy     115 atgcgcccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttt 174
             |||| |  |    | | || ||   | |  ||| ||  |  || ||||   |
Db    4004 cgccgcctcggggatgacgccatggcccgggcggccctcgcatcagtccaacgccctcgc 4063

Qy     175 gaagacaaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgcc 234
            ||| |    |  ||    | ||   |   |  ||  | |   | |||||  |   ||  |
Db    4064 aaaggcccttacaatatcagggtatggaacatggccgcaggcgctggcaagactacccgc 4123

Qy     235 atccatcgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgac 294
           ||||     | |   | ||||| |   |  |  |||       ||  ||    |   ||
Db    4124 atcctcgctgccttcacgcgcgaagacctttacgtctgccccaccaatgcgctcctgcac 4183

Qy     295 gaaatcgagttcgaacgctacgcgcccga 323
           || ||| ||  | ||| |  ||||| |||
Db    4184 gagatccaggccaaactccgcgcgcgcga 4212


RESULT  36
AAA08043
ID   AAA08043 standard; cDNA; 9759 BP.
XX
AC   AAA08043;
XX
DT   19-JUN-2000  (first entry)
XX
DE   Rubella virus cDNA clone sequence SEQ ID NO:1.
XX
KW   Rubella virus; infection; infectious; chimeric construct; anti-viral;
KW   vaccine; birth defect; autoimmune disease; ss.
```

XX
OS Rubella virus.
XX
PN US6054573-A.
XX
PD 25-APR-2000.
XX
PF 02-SEP-1997; 97US-0999733.
XX
PR 28-JUN-1991; 91US-0722334.
PR 19-JUL-1993; 93US-0093453.
PR 02-JUN-1995; 95US-0459041.
XX
PA (UYGE-) UNIV GEORGIA STATE.
XX
PI Abernathy ES, Pougatchev K, Frey TK;
XX
DR WPI; 2000-328366/28.
XX
PT Highly infectious rubella virus clones useful for developing a rubella
PT vaccine that can be safely administered to pregnant and older women
PT without risk of birth defects and autoimmune disease -
XX
PS Disclosure; Column 9-18; 17pp; English.
XX
CC The present invention describes nucleic acid molecules (AAA08044 and
CC AAA08045) which are fragments of the f-Therien rubella virus genome.
CC AAA08044 and AAA08045 are used to replace the corresponding fragments of
CC infectious rubella virus cDNA clone with low specific infectivity
CC referred to as Robo102 to create a chimeric construct with high
CC specific infectivity, Robo302. The highly infectious rubella virus
CC clones are useful as molecular biology tools for studying rubella virus
CC and can be used for developing recombinant vaccines against rubella.
CC The rubella vaccines developed can be safely administered to pregnant
CC and older women without risk of birth defects or autoimmune disease.
CC The present sequence represents a rubella virus cDNA clone used in
CC the exemplification of the present invention.
XX
SQ Sequence 9759 BP; 1458 A; 3784 C; 3007 G; 1510 T; 0 other;

Query Match 2.6%; Score 35.6; DB 21; Length 9759;
Best Local Similarity 45.4%; Pred. No. 6.2;
Matches 122; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

Qy 55 gtcatttatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggt 114
Db 3944 ggcacctgtgccgcgaccgacgaggggctggcccaggcgtactacgacgacctcgaggtg 4003

Qy 115 atgcgccccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgttt 174
Db 4004 cgccgcctcggggatgacgccatggcccgggcggccctcgcatcagtccaacgccctcgc 4063

Qy 175 gaagacaaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgcc 234
Db 4064 aaaggcccttacaatatcagggtatggaacatggccgcaggcgctggcaagactacccgc 4123

Qy 235 atccatcgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgac 294
Db 4124 atcctcgctgccttcacgcgcgaagacctttacgtctgccccaccaatgcgctcctgcac 4183

Qy 295 gaaatcgagttcgaacgctacgcgcccga 323
Db 4184 gagatccaggccaaactccgcgcgcgcga 4212

RESULT 37
AAS54113
ID AAS54113 standard; DNA; 2727 BP.
XX
AC AAS54113;
XX
DT 13-FEB-2002 (first entry)

XX
DE Pseudomonas aeruginosa DNA for cellular proliferation protein #244.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS Pseudomonas aeruginosa.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR P-PSDB; AAU36254.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27; Seq ID No 7750; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2727 BP; 450 A; 1007 C; 854 G; 416 T; 0 other;

Query Match 2.6%; Score 35.4; DB 23; Length 2727;
Best Local Similarity 48.9%; Pred. No. 3.8;
Matches 93; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

Qy 1111 ggcgaccgcgccatggtgccgattggtacttacgagcgcgtaatgccgctagacatcctg 1170
Db 646 gccgacctgcacctggcgctgctgcccggcaccgacgtcgcactgttccacggcatcctg 705

Qy 1171 cctaccctgcttttgcgcgatttaatcgtcggcgataccgacagcgcgcaagcattgggt 1230
Db 706 catatcctgctctgggaggactggatcgaccgttcgttcatcgccgagcataccgagggt 765

Qy 1231 tgcttggaattggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatac 1290
Db 766 ttcgccgacctgaaggaactggtgcgcgactacaccccggccgccgtcgccgacatctgc 825

```
Qy   1291 gaatanggcc 1300
          |     | ||
Db    826 ggcatcgacc 835


RESULT 38
AAS57519
ID   AAS57519 standard; cDNA; 296 BP.
XX
AC   AAS57519;
XX
DT   13-FEB-2002  (first entry)
XX
DE   cDNA #195 encoding portion of a human colon tumour protein.
XX
KW   Human; colon tumour protein; colon cancer; gene therapy; cytostatic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200173027-A2.
XX
PD   04-OCT-2001.
XX
PF   22-MAR-2001; 2001WO-US09246.
XX
PR   24-MAR-2000; 2000US-191597P.
PR   04-MAY-2000; 2000US-202024P.
PR   05-MAY-2000; 2000US-202189P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Meagher MJ,  Xu J,  King GE;
XX
DR   WPI; 2001-611627/70.
XX
PT   New colon tumour proteins and related nucleic acid, useful for
PT   treatment, prevention, diagnosis and monitoring of cancer -
XX
PS   Claim 4; Page 90; 299pp; English.
XX
CC   Th present invention relates to the isolation of novel cDNA sequences
CC   encoding for at least an immunogenic portion of human colon tumour
CC   proteins. The sequences of the invention are useful in pharmaceutical
CC   compositions and vaccines for the prevention and treatment of cancers
CC   such as colon cancer. They are also useful for the diagnosis and
CC   monitoring of such cancers. Antibodies to the colon tumour proteins
CC   and antigen presenting cells that express polynucleotides encoding
CC   colon tumour proteins can be used to inhibit the development of
CC   cancers. T-cells that react specifically with colon tumour proteins
CC   are useful for removing tumour cells from samples (e.g. blood) and
CC   for cancer treatment. The polynucleotides sequences are also useful in
CC   gene therapy. AAS57325-AAS58880 represent the cDNA sequences of the
CC   invention that encode for portions of human colon tumour proteins.
XX
SQ   Sequence 296 BP; 66 A; 86 C; 88 G; 51 T; 5 other;


  Query Match             2.6%;  Score 35;  DB 23;  Length 296;
  Best Local Similarity   57.8%;  Pred. No. 1.7;
  Matches   59;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy   1239 attggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatangg 1298
          |  || ||  || ||    || |   || || |   | ||||||| || ||| | |
Db     74 angggccgccgaggatggggcctgcggctgcctgcccaacccgggcacattcgaggagtg 133

Qy   1299 cccgctgttgcgtaaggtgctggaaaccnttgagaaggaagg 1340
          ||  | |  | | |||| ||||    || || ||| ||| ||
Db    134 ccaccggaagtgcaaggagctgtttcccattcagatggaggg 175


RESULT 39
AAV23811
ID   AAV23811 standard; DNA; 1935 BP.
XX
AC   AAV23811;
XX
DT   29-JUL-1998  (first entry)
XX
DE   Nitrosomonas dnaK gene.
XX
KW   Heatshock promoter; dnaK gene; stress-sensing gene; oxidative stress; ss.
XX
OS   Nitrosomonas europaea.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1935
FT                   /*tag=  a
XX
PN   JP10108678-A.
XX
PD   28-APR-1998.
XX
PF   07-OCT-1996;   96JP-0266320.
XX
PR   07-OCT-1996;   96JP-0266320.
XX
PA   (KURK ) KURITA WATER IND LTD.
XX
DR   WPI; 1998-304973/27.
DR   P-PSDB; AAW54020.
XX
PT   New heat-shock promoter from Nitrosomonas species - useful for, e.g.
PT   measuring oxidative stress caused by ammonia
XX
PS   Example 1; Page 10-13; 18pp; Japanese.
XX
CC   This sequence represents a the dnaK gene of Nitrosomonas
CC   europaea. The heatshock promoter (HP) of the amplified DNA is an example
CC   of the promoter of the invention. The HP can be used in a stress-sensing
CC   gene comprising: (a) a fused DNA fragment comprising HP; and (b) a DNA
CC   fragment positioned downstream of the fragment of (a), comprising a
CC   structural gene encoding a protein for detecting a gene expression. A
CC   microorganism carrying the stress-sensing gene may be used to measure the
CC   oxidative stress caused by ammonia, by culturing it and measuring the
CC   expression of the stress-sensing gene. The method can measure the
CC   oxidative stress easily and rapidly.
XX
SQ   Sequence 1935 BP; 574 A; 442 C; 546 G; 373 T; 0 other;


  Query Match             2.6%;  Score 35;  DB 19;  Length 1935;
  Best Local Similarity   50.3%;  Pred. No. 4.3;
  Matches   77;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy    221 gcaaaatcgccgccatccatcgcggcgaaaagcgcgtacttcagtcggtcgtgattgccg 280
          ||||||| || |    | |||  |||| |       |   |   ||| || |      ||
Db    563 gcaaaattgcggtatacgatctgggcggaggtacgtttgatatttcgatcatcgaaatcg 622

Qy    281 ttgaaggcaacgacgaaatcgagttcgaacgctacgcgcccgaagcgttggcaaacttaa 340
            |||| | | | ||||  | |||||||||      ||  || | |      |   | |
Db    623 cagaagtcgaaggcgaacaccagttcgaagtgctggcaaccaatggtgacactttcctcg 682

Qy    341 gcggcgangaantnngnngcaatctgatccaat 373
          |||| || ||  |      ||   ||||| |||
Db    683 gcggtgaggatttcgactccagagtgatcgaat 715


RESULT 40
AAS86095
ID   AAS86095 standard; cDNA; 323 BP.
XX
AC   AAS86095;
XX
DT   13-FEB-2002  (first entry)
XX
```

DE DNA encoding novel human diagnostic protein #21899.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG21908.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 21899; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 323 BP; 95 A; 86 C; 84 G; 58 T; 0 other;

Query Match 2.6%; Score 34.6; DB 23; Length 323;
Best Local Similarity 58.6%; Pred. No. 2.3;
Matches 58; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 1242 ggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatanggccc 1301
Db 37 ggccgccgaggatggggcctgcggctgcctgcccaacccgggcacattcgaggagtgcca 96

Qy 1302 gctgttgcgtaaggtgctggaaaccnttgagaaggaagg 1340
Db 97 ccggaagtgcaaggagctgtttcccattcagatggaggg 135

RESULT 41
AAF21735
ID AAF21735 standard; DNA; 1139 BP.
XX
AC AAF21735;
XX
DT 27-MAR-2001 (first entry)
XX
DE Human breast and ovarian cancer associated antigen gene SEQ ID 122.
XX
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW nootropic; neurpprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease; ds.
XX
OS Homo sapiens.
XX
PN WO200055173-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05881.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-611515/58.
DR P-PSDB; AAB58832.
XX
PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases -
XX
PS Claim 1; Page 573; 1299pp; English.
XX
CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiant activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and agonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.
XX
SQ Sequence 1139 BP; 217 A; 350 C; 375 G; 188 T; 9 other;

Query Match 2.6%; Score 34.6; DB 21; Length 1139;
Best Local Similarity 58.6%; Pred. No. 4.4;
Matches 58; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 1242 ggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatanggccc 1301
Db 353 ggccgccgaggatggggcctgcggctgcctgcccaacccgggcacattcgaggagtgcca 412

Qy 1302 gctgttgcgtaaggtgctggaaaccnttgagaaggaagg 1340
Db 413 ccggaagtgcaaggagctgtttcccattcagatggaggg 451

```
RESULT  42
AAS86098
ID   AAS86098 standard; cDNA; 1689 BP.
XX
AC   AAS86098;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #21902.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG21911.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 21902; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1689 BP; 348 A; 538 C; 525 G; 278 T; 0 other;


  Query Match             2.6%;  Score 34.6;  DB 23;  Length 1689;
  Best Local Similarity   58.6%;  Pred. No. 5.3;
  Matches   58;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy   1242 ggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatanggccc 1301
           || ||  || ||    || |   || || |   |  |||||||| || |||  |  |||
Db    285 ggccgccgaggatggggcctgcggctgcctgcccaacccgggcacattcgaggagtgcca 344

Qy   1302 gctgttgcgtaaggtgctggaaaccnttgagaaggaagg 1340
```

```
                | |  | |  ||||  ||||    || ||  |||  |||  ||
Db    345 ccggaagtgcaaggagctgtttcccattcagatggaggg 383


RESULT  43
AAS76357
ID   AAS76357 standard; cDNA; 6237 BP.
XX
AC   AAS76357;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #12161.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG12170.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 12161; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 6237 BP; 1645 A; 1625 C; 1779 G; 1187 T; 1 other;


  Query Match             2.6%;  Score 34.6;  DB 23;  Length 6237;
  Best Local Similarity   51.4%;  Pred. No. 10;
  Matches   76;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy     90 gttgcttggcgaagaatatgccggtatgcgccccntgatgaaagtcaaggaaggcgatgc 149
          | ||||   | | |   ||||||   ||  ||  | || ||      | ||||  |||||
```

```
Db   1068 gctgctgtccaaggccaatgccgaggtggcccagtggaggagcaagtacgaagcagatgc 1127

Qy    150 cgtcaaaaaaggccaagtgctgtttgaagacaaaaagnatccgggcgtggtgtttaccgc 209
          | || | |    | | | ||||   || | ||||||  | | ||   ||  | |    |
Db   1128 catccagaggaccgaggagctggaggaggccaaaaaaaagctggcactgcggctgcagga 1187

Qy    210 gccngtttcaggcaaaatcgccgccatc 237
          | | |     |||      || |||| |
Db   1188 ggcagaggagggcgtggaggctgccaac 1215


RESULT 44
AAS73416/c
ID   AAS73416 standard; cDNA; 1050 BP.
XX
AC   AAS73416;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #9220.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG09229.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 9220; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1050 BP; 252 A; 299 C; 264 G; 235 T; 0 other;

  Query Match             2.6%;  Score 34.4;  DB 23;  Length 1050;
  Best Local Similarity   45.1%;  Pred. No. 4.8;
  Matches  110;  Conservative    0;  Mismatches  134;  Indels    0;  Gaps    0;

Qy    128 tgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagacaaaaagn 187
          |||  | |||  |    | || || |   |  |  ||| ||  |  | | | |||| ||
Db    608 TGAGTGACAAAAATCAGGCTGACGCCGCGATCGCGCAACTGACGCATCAGGTCAAACAGT 549

Qy    188 atccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccatcgcggcg 247
          | |    | |  ||  |  || | | |     ||    ||||  |     ||    |||
Db    548 AACTCCACTTCCTGGGTGTCGAGACTGGCGGTGGGTTCATCGAGGATCAGCACTTTGGCA 489

Qy    248 aaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatcgagttcg 307
          | ||  ||     | | |          |   |          ||||  ||  | | | |
Db    488 GAGAGATCGATAGCCCGACAAATAGCGACGATTTGCTGCATCGCGACTGAAAAGCGGTTG 429

Qy    308 aacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngcaatctga 367
          | || | ||||| ||  | |   | || | ||||  || |  |  |  |  ||   ||
Db    428 AGCGGTTCGCGCACGTCGAGGGAGAAACCGTAAGATGCCATCAGTTCGGTGGCGCGCTTT 369

Qy    368 tcca 371
          ||||
Db    368 TCCA 365


RESULT 45
AAS80144/c
ID   AAS80144 standard; cDNA; 1050 BP.
XX
AC   AAS80144;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #15948.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG15957.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 15948; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
```

CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1050 BP; 252 A; 299 C; 264 G; 235 T; 0 other;

Query Match 2.6%; Score 34.4; DB 23; Length 1050;
Best Local Similarity 45.1%; Pred. No. 4.8;
Matches 110; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

```
Qy    128 tgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagacaaaaagn 187
          ||| | |||  |    | || || |   |  |  ||| ||  |  | | | |||| ||
Db    608 TGAGTGACAAAAATCAGGCTGACGCCGCGATCGCGCAACTGACGCATCAGGTCAAACAGT 549

Qy    188 atccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccatcgcggcg 247
          | |   | |  ||  |  || | | |     ||    ||||  |    .||  .|||
Db    548 AACTCCACTTCCTGGGTGTCGAGACTGGCGGTGGGTTCATCGAGGATCAGCACTTTGGCA 489

Qy    248 aaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatcgagttcg 307
          | ||  ||     | | |          |   |         ||||  ||  | | | |
Db    488 GAGAGATCGATAGCCCGACAAATAGCGACGATTTGCTGCATCGCGACTGAAAAGCGGTTG 429

Qy    308 aacgctacgcgcccgaagcgttggcaaacttaagcggcgangaantnngnngcaatctga 367
          | || | ||||| ||  | |   | || | ||||  || |  |  |  |  ||   ||
Db    428 AGCGGTTCGCGCACGTCGAGGGAGAAACCGTAAGATGCCATCAGTTCGGTGGCGCGCTTT 369

Qy    368 tcca 371
          ||||
Db    368 TCCA 365
```

Search completed: June 30, 2002, 09:49:45
Job time: 53218 sec